

F:151-163/Product: neurotensin #status experimental <NTS>  
F:151/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status experime

Query Match 89.2%; Score 794; DB 1; Length 170;  
Best Local Similarity 90.0%; Pred. No. 9.6e-65;  
Matches 153; Conservative 3; Mismatches 14; Indels 0; Gaps 0;

Qy 1 MMAGMKIQLVCMILLAFSSWSLCSSEEMKALEADFLTNMHTSKISKAHVPSWKMTLLN 60  
Db 1 MMAGMKIQLVCMILLAFSSWSLCSSEEMKALEADFLTNMHTSKISKASVSSWKMTLLN 60  
Qy 61 VCSLVNNLNSPAETGEVHEELVARRKLPALDGFSLKAMLTIVQLHKICHSAFOHWE 120  
Db 61 VCSFVNNLSQAETGEVHEELVARRKLPALDGFSLKAMLTIVQLKICHSAFOHWE 120  
Qy 121 LIQEDILDTGNDKNGKEEVIRKIPYILKROLYENKPRPYILKRDSYYY 170  
Db 121 LIQEDVLDAGNDKNEKEEVIRKIPYILKROLYENKPRPYILKRGSYYY 170

## RESULT 2

A28146  
neurotensin / neurotensin N precursor - rat  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 30-Sep-1989 #sequence\_revision 30-Sep-1989 #text\_change 13-Mar-1997  
C:Accession: A28146; A56789  
R:Kislauskis, E.; Bullock, B.; McNeil, S.; Dobner, P.R.  
J. Biol. Chem. 263, 4963-4968, 1988  
A:Title: The rat gene encoding neurotensin and neurotensin N. Structure, tissue-specific  
A:Reference number: A28146; MUID:88169625; PMID:2832414  
A:Accession: A28146  
A:Molecule type: mRNA  
A:Residues: 1-169 <KIS>  
R:Bidard, J.N.; de Nadaï, F.; Rovere, C.; Moinier, D.; Laur, J.; Martinez, J.; Cuber, J.  
Biochem. J. 291, 225-233, 1993  
A:Title: Immunological and biochemical characterization of processing products from the  
A:Reference number: A56789; MUID:93228610; PMID:8471039  
A:Accession: A56789  
A:Status: preliminary  
A:Molecule type: protein  
A:Residues: 23-42 <RID>  
A:Experimental source: medullary thyroid carcinoma 6-23 cell extracts  
A:Note: sequence extracted from NCBI backbone (NCBIP:129397)  
C:Superfamily: neurotensin  
C:Keywords: neuropeptide

Query Match 79.9%; Score 711; DB 2; Length 169;  
Best Local Similarity 78.1%; Pred. No. 3.2e-57;  
Matches 132; Conservative 17; Mismatches 20; Indels 0; Gaps 0;

Qy 2 MAGMKIQLVCMILLAFSSWSLCSSEEMKALEADFLTNMHTSKISKAHVPSWKMTLLN 61  
Db 1 MIGNVLQVCLTLAFSSWSLCSSEEDVRALEADLLTNMHSKSVSGSPPSWKMTLLN 60  
Qy 62 CSLVNNLNSPAETGEVHEELVARRKLPALDGFSLKAMLTIVQLHKICHSAFOHWE 121  
Db 61 CSLINNLSAAEAGEMDDDLVAKRLPLVLDLDFSLALLTVFQLKICHSAFOHWEI 120  
Qy 122 IQEDILDTGNDKNGKEEVIRKIPYILKROLYENKPRPYILKRDSYYY 170  
Db 121 IQEDILDHGNEKTEKEEVIRKIPYILKROLYENKPRPYILKRASYYY 169

## RESULT 3

158190  
neurotensin/neurotensin N - human (fragment)  
C:Species: Homo sapiens (man)  
C:Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 16-Jul-1999  
C:Accession: 158190  
R:Bean, A.J.; Bagerlind, A.; Hokfelt, T.; Dobner, P.R.  
Neuroscience 50, 259-268, 1992  
A:Title: Cloning of human neurotensin/neurotensin N genomic sequences and expression in t  
A:Reference number: 158190; MUID:93063858; PMID:1436492

A:Accession: I58190

A:Status: preliminary; translated from GB/EMBL/DBDJB

A:Molecule type: DNA

A:Residues: 1-120 <RES>

A:Cross-references: GB:547339; NID:g258822; PIDN:AAB23934.1; PID:g258823

C:Superfamily: neurotensin

Query Match 70.3%; Score 626; DB 2; Length 120;  
Best Local Similarity 100.0%; Pred. No. 1.1e-49;  
Matches 120; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MMAGMKIQLVCMILLAFSSWSLCSSEEMKALEADFLTNMHTSKISKAHVPSWKMTLLN 60  
Db 1 MMAGMKIQLVCMILLAFSSWSLCSSEEMKALEADFLTNMHTSKISKAHVPSWKMTLLN 60  
Qy 61 VCSLVNNLNSPAETGEVHEELVARRKLPALDGFSLKAMLTIVQLHKICHSAFOHWE 120  
Db 61 VCSLVNNLNSPAETGEVHEELVARRKLPALDGFSLKAMLTIVQLHKICHSAFOHWE 120

## RESULT 4

H84830  
hypothetical protein At2g40550 [imported] - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 02-Feb-2001  
C:Accession: H84830  
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; N.  
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Unayam, L.; Tallon, L.  
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.B.; Fraser, C.M.; Venter, J.  
Nature 402, 761-768, 1999  
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
A:Reference number: A84420; MUID:20083487; PMID:10617197  
A:Accession: H84830  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-550 <STO>  
A:Cross-references: GB:AE002093; NID:g2651302; PIDN:AAB87582.1; GSPDB:GN00139

C:Genetics:

A:Gene: At2g40550

A:Map position: 2

Query Match 10.1%; Score 90; DB 2; Length 550;  
Best Local Similarity 22.2%; Pred. No. 2.5;  
Matches 24; Conservative 28; Mismatches 44; Indels 12; Gaps 3;

Qy 26 SEEMKALEADFLTNMHTSKISKAHV-----PSWKMTLLNVCSLVNNLNSPAETGEVH 79  
Db 436 SEGKSNIMPADLVLPQPSQVNSLEVITPETAETWRCYLATCKSLSHSI---GQELQQV 492  
Qy 80 EELVARRKLPALDGFSLKAMLTIVQLHKICHSR---AFQHWELIOE 124  
Db 493 ENDLVAARQTRSLGSDQLSRLLTMARMSVSYGETTSLSEHWQNVLE 540

## RESULT 5

H70183  
conserved hypothetical protein BB0673 - Lyme disease spirochete  
C:Species: Borrelia burgdorferi (Lyme disease spirochete)  
C:Date: 13-Feb-1998 #sequence\_revision 13-Feb-1998 #text\_change 08-Oct-1999  
C:Accession: H70183  
R:Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; White,  
son, D.; Peterson, J.; Kerlavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; Vugt,  
; Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B.  
Nature 390, 580-586, 1997  
A:Authors: Smith, H.O.; Venter, J.C.  
A:Title: Genomic sequence of a Lyme disease spirochaete, Borrelia burgdorferi.  
A:Reference number: A70100; MUID:98065943; PMID:9403685  
A:Accession: H70183  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-171 <KLE>  
A:Cross-references: GB:AE001168; GB:AE000783; NID:g2688598; PIDN:AAC67020.1; PID:g268860;  
A:Experimental source: strain B31

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 29, 2003, 17:15:54 ; Search time 21 seconds  
(without alignments)

778.508 Million cell updates/sec

Title: US-09-927-565A-1

Perfect score: 890

Sequence: 1 MMAGMKIQLVCMILLAFSSW.....QLYENKRRPYILKRDVYY 170

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 76:\*

1: pir1.\*

2: pir2.\*

3: pir3.\*

4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	794	89.2	170	1	neurotensin precu
2	711	79.9	169	2	neurotensin / neur
3	626	70.3	120	2	neurotensin/neurom
4	90	10.1	550	2	hypothetical prote
5	87.5	9.8	171	2	conserved hypothet
6	86.5	9.7	499	2	conserved hypothet
7	86.5	9.7	605	2	hypothetical prote
8	86.5	9.7	610	2	hypothetical prote
9	84.5	9.5	1269	2	protein W09G10.4 (
10	81.5	9.2	230	1	H+-transporting tw
11	79.5	8.9	479	2	CIL protein - vari
12	79.5	8.9	479	2	hypothetical prote
13	79.5	8.9	479	2	hypothetical prote
14	79.5	8.9	479	2	EIL protein - vari
15	79.5	8.9	651	2	hypothetical prote
16	79	8.9	811	2	hypothetical prote
17	79	8.9	819	1	endopeptidase La (
18	79	8.9	819	2	Lon ATP-dependent
19	79	8.9	1626	2	DNA topoisomerase
20	78.5	8.8	479	2	EIL protein - vacc
21	78	8.8	283	2	probable lipoprote
22	78	8.8	537	2	acrosomal protein
23	78	8.8	837	2	hypothetical prote
24	78	8.8	908	2	disease resistance
25	77	8.7	1252	2	probable type II D
26	76.5	8.6	273	2	hypothetical prote
27	76.5	8.6	377	2	hypothetical prote
28	76.5	8.6	633	2	probable mureinpep
29	76	8.5	1857	2	hypothetical prote

## ALIGNMENTS

### RESULT 1

UNDG

neurotensin precursor - dog

N;Contains: neuromedin N; neuromedin N-125; neurotensin

C;Species: Canis lupus familiaris (dog)

C;Date: 31-Dec-1992 #sequence revision 31-Dec-1992 #text change 18-Jun-1999

C;Accession: A28025; B60319; A60324; JN0293; A36272

R;Dobner, P.R.; Barber, D.L.; Villia-Komaroff, L.; McKiernan, C.

Proc. Natl. Acad. Sci. U.S.A. 84, 3516-3520, 1987

A;Title: Cloning and sequence analysis of cDNA for the canine neurotensin/neuromedin N p

A;Reference number: A28025; MUID:87204168; PMID:3472221

A;Accession: A28025

A;Molecule type: mRNA

A;Residues: 1-170 <DOB>

A;Cross-references: GB:M16443; NID:G164033; PIDN:AAA30878.1; PID:G164034

R;Mittra, S.P.; Muraki, K.; Brown, D.R.; Parsons, A.M.; Carraway, R.E.

Regul. Pept. 28, 11-22, 1990

A;Title: Canine neurotensin, neurotensin(6-13) and neuromedin N: primary structures and

A;Reference number: A60319; MUID:90222594; PMID:2158127

A;Accession: B60319

A;Molecule type: protein

A;Residues: 143-148 <MIT>

A;Accession: A60319

A;Molecule type: protein

A;Residues: 152-163 <MIT2>

R;Mogard, M.H.; Reeve Jr., J.R.; Shively, J.E.; Ben-Avram, C.M.; Eysselein, V.E.; Walsh,

Regul. Pept. 14, 313-321, 1986

A;Title: Isolation and characterization of a neurotensin-like decapeptide from a canine

A;Reference number: A60324; MUID:86314883; PMID:3749527

A;Accession: A60324

A;Molecule type: protein

A;Residues: 154-163 <MOG>

R;Carraway, R.E.; Mittra, S.P.

Biochem. Biophys. Res. Commun. 179, 301-308, 1991

A;Title: Purification of large neuromedin N (NMN) from canine intestine and its identific

A;Reference number: JN0293; MUID:91354266; PMID:1883359

A;Accession: JN0293

A;Molecule type: protein

A;Residues: 24-43 <CAR>

A;Experimental source: small intestine

R;Carraway, R.E.; Mittra, S.P.

J. Biol. Chem. 265, 8627-8631, 1990

A;Title: Differential processing of neurotensin/neuromedin N precursor(s) in canine brain

A;Reference number: A36272; MUID:90256783; PMID:2341398

A;Accession: A36272

A;Molecule type: protein

A;Residues: 128-147 <CA2>

C;Superfamily: neurotensin

C;Keywords: hormone; neuropeptide; pyroglutamic acid

F;1-23/Domain: signal sequence #status predicted <SIG>

F;24-148/Product: large neuromedin N-125 #status experimental <LNM>

F;143-148/Product: neuromedin N #status experimental <NMN>

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 29, 2003, 17:14:33 ; Search time 17 Seconds  
(without alignments)  
470.267 Million cell updates/sec

Title: US-09-927-565A-1

Perfect score: 890

Sequence: 1 MMAGMKIQLVCMLLAFSSW.....QLYENKPRRPYILKRDSYYY 170

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_41.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	890	100.0	170	1	NEUT_HUMAN
2	804	90.3	169	1	NEUT_BOVIN
3	794	89.2	170	1	NEUT_CANFA
4	742	83.4	169	1	NEUT_MOUSE
5	711	79.9	169	1	NEUT_RAT
6	83.5	9.4	832	1	EP2_CRYPV
7	81.5	9.2	102	1	R510_SULTO
8	81.5	9.2	230	1	NTPD_ENTHR
9	79.5	8.9	196	1	ADEN_ADEBA
10	79.5	8.9	346	1	XYNB_THENE
11	79.5	8.9	479	1	PAP1_VACCV
12	79.5	8.9	479	1	PAP1_VARV
13	79	8.9	811	1	RPOP_NEUN
14	79	8.9	819	1	LON_CHLNP
15	79	8.9	1626	1	TP2B_HUMAN
16	78.5	8.8	416	1	PALY_VITVI
17	78.5	8.8	429	1	CGBI_MESAU
18	78.5	8.8	479	1	PAP1_VACCC
19	78	8.8	283	1	Y43C_MYCPN
20	78	8.8	908	1	RP8B_ARATH
21	76.5	8.6	377	1	Y835_METJA
22	75.5	8.5	401	1	PILC_PSEPU
23	75	8.4	505	1	TCMO_ZINEL
24	75	8.4	506	1	TCMO_SOYBN
25	75	8.4	1612	1	TP2B_CRILLO
26	74	8.3	585	1	Y187_MYCGE
27	74	8.3	1067	1	MICA_HUMAN
28	74	8.3	1612	1	TP2B_MOUSE
29	73.5	8.3	3680	1	DND_CANFA
30	73	8.2	843	1	CYAA_HAEIN
31	73	8.2	908	1	R8L4_ARATH
32	73	8.2	1939	1	MYH6_HUMAN
33	73	8.2	3433	1	UTRO_HUMAN

#### RESULT 1

NEUT\_HUMAN  
ID NEUT\_HUMAN STANDARD; PRT; 170 AA.  
AC P30990;  
DT 01-JUL-1993 (Rel. 26, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Neurotensin/neuromedin N precursor [Contains: Large neuromedin N (NmN-125); Neuromedin N (NmN) (NN); Neurotensin (NT); Tail peptide].  
GN NTS.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=98191186; PubMed=9530155;  
RA Dong Z., Wang X., Zhao Q., Townsend C.M. Jr., Evers B.M.;  
RT "DNA methylation contributes to expression of the human  
neurotensin/neuromedin N gene."  
RL Am. J. Physiol. 274:G535-G543(1998).  
RN [2]  
RP SEQUENCE OF 1-120 FROM N.A.  
RX MEDLINE=93063858; PubMed=1436492;  
RA Bean A.J., Pagerlind A., Hoekfelt T., Dobner P.R.;  
RT "Cloning of human neurotensin/neuromedin N genomic sequences and  
expression in the ventral mesencephalon of schizophrenics and age/sex  
matched controls."  
RL Neuroscience 50:259-268(1992).  
CC  
CC -!- FUNCTION: Neurotensin may play an endocrine or paracrine role  
in the regulation of fat metabolism. It causes contraction of  
smooth muscle.  
CC  
CC -!- SUBCELLULAR LOCATION: Secreted; Packaged within secretory  
vesicles.  
CC  
CC -!- SIMILARITY: BELONGS TO THE NEUROTENSIN FAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; U91618; AAB50564.1; -;  
DR EMBL; S47339; AAB23934.1; -;  
DR PIR; I58190; I58190.  
DR Genew; HGNC:8038; NTS.  
DR MIM; 162650; -;  
GO; GO:0005625; C:soluble fraction; TAS.  
KW Cleavage on pair of basic residues; Vasoactive; Signal.  
FT SIGNAL 1 23 POTENTIAL.  
FT CHAIN 24 148 LARGE NEUROMEDIN N (NMN-125).  
FT PEPTIDE 144 148 NEUROMEDIN N.  
FT PEPTIDE 151 163 NEUROTENSIN.  
FT PEPTIDE 166 170 TAIL PEPTIDE (POTENTIAL).

#### ALIGNMENTS

34	72	8.1	233	1	BCLX_HUMAN	Q07817	homo sapien
35	72	8.1	465	1	MYH6_RABIT	P04460	oryctolagus
36	72	8.1	505	1	TCMO_ARATH	P92994	arabidopsis
37	72	8.1	505	1	TCMO_CICAR	O81928	cicer ariet
38	72	8.1	879	1	YN65_YEAST	P42837	saccharomyc
39	72	8.1	1536	1	SIN3_YEAST	P22579	saccharomyc
40	72	8.1	2269	1	WDR9_HUMAN	Q9NS16	homo sapien
41	71.5	8.0	682	1	PAL1_POPKI	P45731	populus kit
42	71	8.0	301	1	B3G5_PANCA	Q9N294	p beta-1.3-
43	71	8.0	325	1	THIL_BACSU	O05514	bacillus su
44	71	8.0	445	1	XAV5_XANCV	P19520	xanthomonas
45	71	8.0	505	1	TCMO_POPTM	O24312	populus tre

SQ SEQUENCE 170 AA; 19795 MW; 66E2E146DA08E3C7 CRC64;

Query Match 100.0%; Score 890; DB 1; Length 170;  
 Best Local Similarity 100.0%; Pred. No. 1.7e-74;  
 Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MMAGKIQVCMVLLAFSSWSLCSSEEMKALEADFLTNMHTSKISKAHVPSWKMTLLN 60  
 DB 1 MMAGKIQVCMVLLAFSSWSLCSSEEMKALEADFLTNMHTSKISKAHVPSWKMTLLN 60  
 QY 61 VCSLVNNLNSPAEETGEVHEELVARRKLPALDGFSLSEAMLTIIYQLHKICHSAFQHW 120  
 DB 61 VCSLVNNLNSPAEETGEVHEELVARRKLPALDGFSLSEAMLTIIYQLHKICHSAFQHW 120  
 QY 121 LIQEDILDTGNDKNGKEVIRKPIYILKQLYENKPRRPYILKRDSDYY 170  
 DB 121 LIQEDILDTGNDKNGKEVIRKPIYILKQLYENKPRRPYILKRDSDYY 170

RESULT 2

NEUT\_BOVIN STANDARD; PRT; 169 AA.

AC P01156;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 01-JUL-1989 (Rel. 11, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Neurotensin/neuromedin N precursor [Contains: Large neuromedin N (NmN-125); Neuromedin N (NmN) (NN); Neurotensin (NT); Tail peptide].  
 OS Bos taurus (Bovine).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovidae; Bovinae; Bos.  
 OX NCBI\_TaxID=9913;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=88169625; PubMed=2832414;  
 RA Kislaukis E., Bullock B., McNeil S., Dobner P.R.;  
 RT "The rat gene encoding neurotensin and neuromedin N. Structure, tissue-specific expression, and evolution of exon sequences.";  
 RL J. Biol. Chem. 263:4963-4968(1988).  
 RN [2]  
 RP SEQUENCE OF 150-162.  
 RC TISSUE=Hypothalamus;  
 RX MEDLINE=75095679; PubMed=1167549;  
 RA Caraway R., Leeman S.E.;  
 RT "The amino acid sequence of a hypothalamic peptide, neurotensin.";  
 RL J. Biol. Chem. 250:1907-1911(1975).  
 RN [3]  
 RP SYNTHESIS OF NEUROTENSIN.  
 RX MEDLINE=75095679; PubMed=1112838;  
 RA Caraway R., Leeman S.E.;  
 RT "The synthesis of neurotensin.";  
 RL J. Biol. Chem. 250:1912-1918(1975).  
 CC -!- FUNCTION: Neurotensin may play an endocrine or paracrine role in the regulation of fat metabolism. It causes contraction of smooth muscle.  
 CC -!- SUBCELLULAR LOCATION: Secreted; Packaged within secretory vesicles.  
 CC -!- TISSUE SPECIFICITY: Brain and gut.  
 CC -!- SIMILARITY: BELONGS TO THE NEUROTENSIN FAMILY.

-----  
 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----  
 DR EMBL; M18621; AAA30668.1; "  
 KW Cleavage on pair of basic residues; Vasoactive; Signal;  
 KW Pyrrolidone carboxylic acid.

FT SIGNAL 1 22 POTENTIAL.  
 FT CHAIN 23 147 LARGE NEUROMEDIN N.  
 FT PEPTIDE 142 147 NEUROMEDIN N.  
 FT PEPTIDE 150 162 NEUROTENSIN.  
 FT PEPTIDE 165 169 TAIL PEPTIDE (POTENTIAL).  
 FT MOD RES 150 150 PYRROLIDONE CARBOXYLIC ACID.  
 SQ SEQUENCE 169 AA; 19712 MW; 7B78760D5E4D7D32 CRC64;

Query Match 90.3%; Score 804; DB 1; Length 169;  
 Best Local Similarity 90.5%; Pred. No. 1.2e-66;  
 Matches 153; Conservative 4; Mismatches 12; Indels 0; Gaps 0;

QY 2 MAGKIQVCMVLLAFSSWSLCSSEEMKALEADFLTNMHTSKISKAHVPSWKMTLLN 61  
 DB 1 MAGKIQVCMVLLAFSSWSLCSSEEMKALEADFLTNMHTSKISKAHVPSWKMTLLN 60  
 QY 62 CSLVNNLNSPAEETGEVHEELVARRKLPALDGFSLSEAMLTIIYQLHKICHSAFQHW 121  
 DB 61 CSLVNNLNSPAEETGEVHEELVARRKLPALDGFSLSEAMLTIIYQLHKICHSAFQHW 120  
 QY 122 IQEDILDTGNDKNGKEVIRKPIYILKQLYENKPRRPYILKRDSDYY 170  
 DB 121 IQEDILDTGNDKNGKEVIRKPIYILKQLYENKPRRPYILKRDSDYY 169

RESULT 3

NEUT\_CANFA STANDARD; PRT; 170 AA.

AC P10673; O9TS28;  
 DT 01-JUL-1989 (Rel. 11, Created)  
 DT 01-JUL-1989 (Rel. 11, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Neurotensin/neuromedin N precursor [Contains: Large neuromedin N (NmN-125); Neuromedin N (NmN) (NN); Neurotensin (NT); NT-tail; Tail peptide].  
 OS Canis familiaris (Dog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.  
 OX NCBI\_TaxID=9615;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=87204168; PubMed=3472221;  
 RA Dobner P.R., Barber D.L., Villa-Komaroff L., McKiernan C.;  
 RT "Cloning and sequence analysis of cDNA for the canine neurotensin/neuromedin N precursor.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 84:3516-3520(1987).  
 RN [2]  
 RP SEQUENCE OF 128-148.  
 RX MEDLINE=90256783; PubMed=2341398;  
 RA Caraway R.E., Mitra S.P.;  
 RT "Differential processing of neurotensin/neuromedin N precursor in canine brain and intestine.";  
 RL J. Biol. Chem. 265:8627-8631(1990).  
 RN [3]  
 RP SEQUENCE OF 24-43.  
 RC TISSUE=Intestine;  
 RX MEDLINE=91354266; PubMed=1883359;  
 RA Caraway R.E., Mitra S.P.;  
 RT "Purification of large neuromedin N (NMN) from canine intestine and its identification as NMN-125.";  
 RL Biochem. Biophys. Res. Commun. 179:301-308(1991).  
 RN [4]  
 RP SEQUENCE OF 151-170.  
 RC TISSUE=Intestine;  
 RX MEDLINE=93157144; PubMed=1494486;  
 RA Caraway R.E., Mitra S.P., Salmons R.;  
 RT "Isolation and quantitation of several new peptides from the canine neurotensin/neuromedin N precursor.";  
 RL Peptides 13:1039-1047(1992).  
 CC -!- FUNCTION: Neurotensin may play an endocrine or paracrine role in the regulation of fat metabolism. It causes contraction of smooth muscle.

CC -!- SUBCELLULAR LOCATION: Secreted; Packaged within secretory vesicles.  
CC -!- SIMILARITY: BELONGS TO THE NEUROTENSIN FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; M16443; AAA30878.1; -;  
CC PIR; A28025; UNDC.  
CC KW Pyroglutamate carboxylic acid.  
CC FT SIGNAL 1 23  
CC FT CHAIN 24 148 LARG NEUROMEDIN N.  
CC FT PEPTIDE 143 148 NEUROMEDIN N.  
CC FT PEPTIDE 151 163 NEUROTENSIN.  
CC FT PEPTIDE 151 170 NT-TAIL.  
CC FT PEPTIDE 166 170 TAIL PEPTIDE (POTENTIAL).  
CC FT MOD RES 151 151 PYROGLUTAMATE CARBOXYLIC ACID.  
CC SQ SEQUENCE 170 AA; 19863 MW; A54700163AC54962 CRC64;  
CC -----  
CC Query Match 89.2%; Score 794; DB 1; Length 170;  
CC Best Local Similarity 90.0%; Pred. No. 1e-65;  
CC Matches 153; Conservative 3; Mismatches 14; Indels 0; Gaps 0;  
CC -----  
CC Qy 1 MMAGKIQLVCMILLAFSSWSLCSDEEMKALEADFLTNMHTSKISKAHVPSWMTLLN 60  
CC Db 1 MMAGKIQLVCMILLAFSSWSLCSDEEMKALEADFLTNMHTSKISKASVSSWMTLLN 60  
CC Qy 61 VCSLVNNLSPAEETGEVHEBELVARRKLPALDGFSLSEAMLTIVQLHKICHSAFOHWE 120  
CC Db 61 VCSFVNNSQAEETGEVHEBELVARRKLPALDGFSLSEAMLTIVQLHKICHSAFOHWE 120  
CC Qy 121 LIQEDILDTGNDKNGKEVIRKIPYILKRLQYENKPRPYILKRDYSYY 170  
CC Db 121 LIQEDVLDAGNDKNGKEVIRKIPYILKRLQYENKPRPYILKRGYSYY 170  
CC -----  
CC RESULT 4  
CC NEUT MOUSE  
CC ID NEUT MOUSE STANDARD; PRT; 169 AA.  
CC AC Q9D3E9;  
CC DT 28-FEB-2003 (Rel. 41, Created)  
CC DT 28-FEB-2003 (Rel. 41, Last sequence update)  
CC DT 28-FEB-2003 (Rel. 41, Last annotation update)  
CC DE Neurotensin/neuromedin N precursor [Contains: Large neuromedin N (NmN-  
CC DE 125); Neuromedin N (NmN) (NN); Neurotensin (NT); Tail peptide].  
CC GN NTS.  
CC OS Mus musculus (Mouse).  
CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
CC OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
CC OX NCBI\_TaxID=10090;  
CC RN [1]  
CC RN SEQUENCE FROM N.A.  
CC RP STRAIN=129;  
CC RA Dobner P.R., Fadel J., Deitemeyer N., Carraway R.E., Deutch A.Y.;  
CC RT "Neurotensin-deficient mice show altered responses to antipsychotic  
CC RT drugs";  
CC RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.  
CC RN [2]  
CC RP SEQUENCE FROM N.A.  
CC RA Ahn H.-J., Cho J.-J.;  
CC RT "Mouse proneurotensin/proneuromedin N is induced in mast cell line  
CC RT after IGE cross-linking";  
CC RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.  
CC RN [3]  
CC RP SEQUENCE FROM N.A.  
CC RC STRAIN=C57BL/6J; TISSUE=Intestine;  
CC RC MEDLINE=21085660; PubMed=11217851;  
CC -----  
CC Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
CC Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Yamanaka S.,  
CC Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,  
CC Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,  
CC Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
CC Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,  
CC Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,  
CC Schriml L.M., Staehli F., Suzuki R., Tomita M., Wagner L., Washio T.,  
CC Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
CC Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
CC Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
CC Guttinger S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
CC Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaeerts P.,  
CC Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
CC Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-P.,  
CC Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,  
CC Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohetsuki S.,  
CC Hayashizaki Y.;  
CC "Functional annotation of a full-length mouse cDNA collection.";  
CC Nature 409:685-690(2001).  
CC -!- FUNCTION: Neurotensin may play an endocrine or paracrine role  
CC in the regulation of fat metabolism. It causes contraction of  
CC smooth muscle (By similarity).  
CC -!- SUBCELLULAR LOCATION: Secreted; packaged within secretory  
CC vesicles (By similarity).  
CC -!- SIMILARITY: BELONGS TO THE NEUROTENSIN FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; AF348489; AAK28626.1; -;  
CC EMBL; AF304160; AAK15263.1; -;  
CC EMBL; AK017212; BAB30636.1; -;  
CC MGD; MGI:1914655; 5033428B16Rik.  
CC KW Cleavage on pair of basic residues; Vasoactive; Signal.  
CC FT SIGNAL 1 22  
CC FT CHAIN 23 147 LARG NEUROMEDIN N.  
CC FT PEPTIDE 142 147 NEUROMEDIN N.  
CC FT PEPTIDE 150 162 NEUROTENSIN.  
CC FT PEPTIDE 165 169 TAIL PEPTIDE (POTENTIAL).  
CC SQ SEQUENCE 169 AA; 19607 MW; D153319F5F0925DC CRC64;  
CC -----  
CC Query Match 83.4%; Score 742; DB 1; Length 169;  
CC Best Local Similarity 81.7%; Pred. No. 5.6e-61;  
CC Matches 138; Conservative 15; Mismatches 16; Indels 0; Gaps 0;  
CC -----  
CC Qy 2 MMAGKIQLVCMILLAFSSWSLCSDEEMKALEADFLTNMHTSKISKAHVPSWMTLLN 61  
CC Db 1 MRGMNLQVLCTLLAFSSWSLCSDEEDVRALEADLLTNMHTSKISKASPPSWMTLLN 60  
CC Qy 62 CSLVNNLSPAEETGEVHEBELVARRKLPALDGFSLSEAMLTIVQLHKICHSAFOHWE 121  
CC Db 61 CSLVNNLSPAEETGEVHEBELVARRKLPALDGFSLSEAMLTIVQLHKICHSAFOHWE 120  
CC Qy 122 IQEDILDTGNDKNGKEVIRKIPYILKRLQYENKPRPYILKRDYSYY 170  
CC Db 121 IQEDILDTGNDKNGKEVIRKIPYILKRLQYENKPRPYILKRGYSYY 169  
CC -----  
CC RESULT 5  
CC NEUT RAT  
CC ID NEUT RAT STANDARD; PRT; 169 AA.  
CC AC P20058; Q9QV80;  
CC DT 01-FEB-1991 (Rel. 17, Created)  
CC DT 01-FEB-1991 (Rel. 17, Last sequence update)  
CC DT 28-FEB-2003 (Rel. 41, Last annotation update)  
CC DE Neurotensin/neuromedin N precursor [Contains: Large neuromedin N (NmN-  
CC DE 125); Neuromedin N (NmN) (NN); Neurotensin (NT); Tail peptide].  
CC -----

NTS.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_taxid=10116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=88169625; PubMed=2832414;  
RA Kiehlaukis E., Bullock B., McNeil S., Dobner P.R.;  
RT "The rat gene encoding neurotensin and neuromedin N. Structure,  
RT tissue-specific expression, and evolution of exon sequences.";  
RL J. Biol. Chem. 263:4963-4968(1988).  
RN [2]  
RP SEQUENCE OF 23-42, AND PROCESSING.  
RX MEDLINE=93228610; PubMed=8471039;  
RA Bidard J.-N., de Nadai F., Rovere C., Moinier D., Laur J.,  
RA Martinez J., Cuber J.-C., Kitabgi P.;  
RT "Immunological and biochemical characterization of processing products  
RT from the neurotensin/neuromedin N precursor in the rat medullary  
RL thymoid carcinoma 6-23 cell line.";  
RL Biochem. J. 291:225-233(1993).  
RN [3]  
RP PROCESSING.  
RX MEDLINE=93215530; PubMed=8462460;  
RA de Nadai F., Rovere C., Bidard J.-N., Laur J., Martinez J.,  
RA Cuber J.-C., Kitabgi P.;  
RT "Biosynthesis and posttranslational processing of the  
RT neurotensin/neuromedin N precursor in the rat medullary thymoid  
RL carcinoma 6-23 cell line. Effect of dexamethasone.";  
RL Endocrinology 132:1614-1620(1993).  
CC -!- FUNCTION: Neurotensin may play an endocrine or paracrine role  
CC in the regulation of fat metabolism. It causes contraction of  
CC smooth muscle.  
CC -!- SUBCELLULAR LOCATION: Secreted; Packaged within secretory  
CC vesicles.  
CC -!- SIMILARITY: BELONGS TO THE NEUROTENSIN FAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; M21187; AAA41712.1; ALT\_INIT.  
DR EMBL; M21218; AAA41712.1; JOINED.  
DR PIR; A28146; A28146.  
KW Cleavage on pair of basic residues; Vasoactive; Signal.  
FT SIGNAL 1 22  
FT CHAIN 23 147 LARGE NEUROMEDIN N.  
FT PEPTIDE 142 147 NEUROMEDIN N.  
FT PEPTIDE 150 162 NEUROTENSIN.  
FT PEPTIDE 165 169 TAIL PEPTIDE (POTENTIAL).  
SQ SEQUENCE 169 AA; 19563 MW; 91CF0CF68C3D3C91 CRC64;  
Query Match 79.9%; Score 711; DB 1; Length 169;  
Best Local Similarity 78.1%; Pred. No. 3.8e-58;  
Matches 132; Conservative 17; Mismatches 20; Indels 0; Gaps 0;  
QY 2 MAGMKIQLVCMLLAFSSWSLCSSEEMKALEADFLTNMHTSKSKAHVPSPKMTLLNV 61  
DB 1 MIGNNLQVLCITLAFSSWSLCSSEEDVRALEADLLTNMASKVSKSPSPKMTLLNV 60  
QY 62 CSLVNNLNSPAETGEVHEELVARRKLPALDGFSLFAMLTIIYQLHKICHRSFAQHWE 121  
DB 61 CSLINNLSAAEEAGEMRDDLVAKRKLPLVLDLDFSLFALLTVFQKICRSFAQHWEI 120  
QY 122 IOEDILDGNGKNGKEEVKIPYILKQLYENKPRPYILKRDSSYY 170  
DB 121 IOEDILDGNGKNGKEEVKIPYILKQLYENKPRPYILKRDSSYY 169

RESULT 6  
EF2\_CRYPV STANDARD; PRT; 832 AA.  
AC Q23716;  
DT 15-JUL-1998 (Rel. 36, Created)  
DT 15-JUL-1998 (Rel. 36, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Elongation factor 2 (EF-2).  
OS Cryptosporidium parvum  
OC Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida;  
OC Cryptosporidiidae; Cryptosporidium.  
OX NCBI\_taxid=5807;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX STRAIN=AUCP-1;  
RX MEDLINE=95356792; PubMed=7630379;  
RA Jones D.E., Tu T.D., Mathur S., Sweeney R.W., Clark D.P.;  
RT "Molecular cloning and characterization of a Cryptosporidium parvum  
RT elongation factor-2 gene.";  
RL Mol. Biochem. Parasitol. 71:143-147(1995).  
CC -!- FUNCTION: THIS PROTEIN PROMOTES THE GTP-DEPENDENT TRANSLLOCATION  
CC OF THE NASCENT PROTEIN CHAIN FROM THE A-SITE TO THE P-SITE OF THE  
CC RIBOSOME (BY SIMILARITY).  
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
CC -!- SIMILARITY: BELONGS TO THE GTP-BINDING ELONGATION FACTOR FAMILY.  
CC EF-G/EF-2 SUBFAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; U21667; AAC46607.1;  
DR InterPro; IPR000795; EF\_GTPbind.  
DR InterPro; IPR000640; EF\_G C.  
DR InterPro; IPR005517; EF\_G\_IV.  
DR InterPro; IPR004161; EFTU\_D2.  
DR InterPro; IPR005225; Small\_GTP.  
DR Pfam; PF00679; EFG\_C; 1.  
DR Pfam; PF03764; EFG\_IV; 1.  
DR Pfam; PF00009; GTP\_EFTU; 1.  
DR Pfam; PF03144; GTP\_EFTU\_D2; 1.  
DR PRINTS; PR00315; ELONGATNFTCT.  
DR TIGRFAMs; TIGR00231; small\_GTP; 1.  
DR PROSITE; PS00301; EFACITOR\_GTP; 1.  
KW Elongation factor; Protein biosynthesis; GTP-binding; Phosphorylation.  
FT NP\_BIND 26 33 GTP (BY SIMILARITY).  
FT NP\_BIND 98 102 GTP (BY SIMILARITY).  
FT NP\_BIND 152 155 GTP (BY SIMILARITY).  
FT MOD\_RES 57 57 PHOSPHORYLATION (BY SIMILARITY).  
FT MOD\_RES 59 59 PHOSPHORYLATION (BY SIMILARITY).  
FT MOD\_RES 689 689 DIPHTHAMIDE (BY SIMILARITY).  
SQ SEQUENCE 832 AA; 92761 MW; 577D2DE3D77E3FA CRC64;  
Query Match 9.4%; Score 83.5; DB 1; Length 832;  
Best Local Similarity 24.7%; Pred. No. 6.4;  
Matches 40; Conservative 19; Mismatches 54; Indels 49; Gaps 9;  
QY 22 LCSDESEEMKALEADFL-TNMHTSKSKAH---VPSMKWT-----LLAVCSLVN 66  
DB 666 LC-----EEDMGIRFNLLDVLHADAIHRGAGQITPTCRRVYAAALFASPLLEPMFLV- 721  
QY 67 NLNSPASETEGEV-----HEE-----ELVARRKLPALDGFSLFAMLTIIYQLH 108  
DB 722 EISAPQEVGGIYATLNQRGHVHEBPKSTPQVEIKAYLPVA-DSFKFTTVLRAATSG 780  
QY 109 KICHSRAFOHWELIQEDILDTGND-----KNGKEEV 139  
DB 781 KAPPCQCVFDHWELINGDPLEKSGKTEELVKAIRRRKNKEI 822

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 29, 2003, 17:14:08 ; Search time 41 Seconds  
(without alignments)

658.135 Million cell updates/sec

Title: US-09-927-565A-1

Perfect score: 890

Sequence: 1 MMAGMKIQLVCMLLAFSSW.....QLYENKPRRPYILKRDSSYY 170

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq 19Jun03.\*

1: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1980.DAT.\*  
2: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1981.DAT.\*  
3: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1982.DAT.\*  
4: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1983.DAT.\*  
5: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1984.DAT.\*  
6: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1985.DAT.\*  
7: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1986.DAT.\*  
8: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1987.DAT.\*  
9: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1988.DAT.\*  
10: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1989.DAT.\*  
11: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1990.DAT.\*  
12: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1991.DAT.\*  
13: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1992.DAT.\*  
14: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1993.DAT.\*  
15: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1994.DAT.\*  
16: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1995.DAT.\*  
17: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1996.DAT.\*  
18: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1997.DAT.\*  
19: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1998.DAT.\*  
20: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT.\*  
21: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA2000.DAT.\*  
22: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA2001.DAT.\*  
23: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA2002.DAT.\*  
24: /SIDSI/gcgdata/geneseq/geneseqp-emb1/AA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	890	100.0	170	22 AAU08929	Human preproneurot
2	890	100.0	170	24 ABUS56410	Lung cancer-associ
3	890	100.0	170	24 ABUS56659	Lung cancer-associ
4	664	74.6	129	21 AAG00256	Human secreted pro
5	90	10.1	279	21 AAB29865	Human secreted pro
6	87.5	9.8	322	20 AAY19902	B. burgdorferi ant
7	80	9.0	424	20 AAY34627	Chlamydia pneumoni
8	80	9.0	841	22 ABG13829	Novel human diagno
9	79.5	8.9	652	18 AAW18010	Plasmodium falcipa

10	79	8.9	400	22	ABG02720	Novel human diagno
11	79	8.9	1139	22	ABG22397	Novel human diagno
12	79	8.9	1621	22	AAW25646	Human protein sequ
13	78.5	8.8	550	20	AAW02299	Computer generated
14	78.5	8.8	550	20	AAW02300	Computer generated
15	78.5	8.8	550	20	AAW02302	Computer generated
16	78.5	8.8	550	20	AAW02303	Computer generated
17	78.5	8.8	550	20	AAW02304	Computer generated
18	78.5	8.8	550	20	AAW02307	Computer generated
19	78.5	8.8	550	20	AAW02290	A (E)-beta-farnese
20	78.5	8.8	550	21	AAW0841	Mint E-b-farnesene
21	77.5	8.7	1726	24	AAO16416	Human nucleic acid
22	77.5	8.7	2816	22	AAU68572	Human novel cytoki
23	76.5	8.6	216	23	ABP53327	Crassula barkleyi
24	76.5	8.6	223	21	AAG08259	Arabidopsis thalia
25	76.5	8.6	247	21	AAG08258	Arabidopsis thalia
26	76.5	8.6	273	21	AAG08257	Arabidopsis thalia
27	76.5	8.6	550	20	AAW02306	Computer generated
28	76	8.5	428	24	ABU11835	Human MDT polypep
29	75.5	8.5	491	24	ABJ25609	Aspergillus fumiga
30	75.5	8.5	554	24	ABJ26209	Aspergillus fumiga
31	75.5	8.5	2485	21	ABJ18172	Plasmodium falcipa
32	75	8.4	212	22	AAG44285	Mutant bcl-Xl prot
33	75	8.4	233	21	AAW83223	Bcl-x polypeptide.
34	75	8.4	233	21	AAW69969	Human Bcl-XL prote
35	75	8.4	233	22	AAG64262	Human Bcl-XL prote
36	75	8.4	233	22	AAB50538	Human Bcl-XL prote
37	75	8.4	233	22	AAB47515	Protein encoded by
38	75	8.4	237	23	ABG78480	Wild type BclXl pr
39	75	8.4	411	22	AAU00219	Bcl-Xl-BTR apoptos
40	75	8.4	419	21	AAG22185	Arabidopsis thalia
41	75	8.4	419	21	AAG48520	Arabidopsis thalia
42	75	8.4	485	22	AAU00222	Lfn-Bcl-XL apoptos
43	75	8.4	565	21	AAG22184	Arabidopsis thalia
44	75	8.4	565	21	AAG48519	Arabidopsis thalia
45	75	8.4	602	21	AAG22183	Arabidopsis thalia

#### ALIGNMENTS

RESULT 1  
AAU08929  
ID AAU08929 standard; Protein, 170 AA.  
XX AAU08929;  
AC AAU08929;  
XX  
DT 18-DEC-2001 (first entry)  
XX Human preproneurotensin/neuromedin N.  
XX Human; neurotensin; neuromedin N; nootropic; neuroprotective;  
KW antiparkinsonian; antidiabetic; anticonvulsant; antidepressant;  
KW antithyroid; cytostatic; antiarteriosclerotic; hepatotrophic;  
KW antiinflammatory; cell proliferative disorder; neurological disorder;  
KW endocrine disorders; gene therapy; antisense therapy; HspN;  
KW Alzheimer's disease; anyotrophic lateral sclerosis; depression;  
KW diabetic neuropathy; epilepsy; Down's syndrome; Huntington's disease;  
KW Parkinson's disease; Cushing's disease; diabetes insipidus;  
KW diabetes mellitus; hypoglycaemia; hypoglycaemia; goitre;  
KW arteriosclerosis; atherosclerosis; hepatitis; cancer; leukaemia;  
KW INCYTE2923961; INCYTE1576389; INCYTE1309070.

OS Homo sapiens.  
XX Key Location/Qualifiers  
FH Peptide 1..142  
FT /label= Prepro\_peptide  
FT Peptide 127..136  
FT /note= Unknown  
FT /claim 1"  
FT Peptide 143..148

FT Peptide /label= Neurotensin  
 FT 151...163  
 XX /label= Neuromedin\_N  
 PN US6274720-B1.  
 PD 14-AUG-2001.  
 XX  
 PF 31-DEC-1997; 97US-0002114.  
 XX  
 PR 31-DEC-1997; 97US-0002114.  
 XX  
 PA (INCY-) INCYTE GENOMICS INC.  
 PI Lal P, Shah P, Corley NC;  
 XX WPI; 2001-588905/66.  
 DR N-PSDB; AAS13802.  
 XX  
 XX Isolated polynucleotide encoding human preneurotensin/neuromedin N  
 PT is used in the diagnosis, treatment and prevention of cell  
 PT proliferative, neurological and endocrine disorders -  
 XX  
 PS Claim 1; Fig 1; 27pp; English.  
 XX  
 CC The invention relates to an isolated polynucleotide sequence encoding  
 CC human neurotensin/neuromedin N (HPPN) neuropeptides. HPPN polynucleotides  
 CC may be used for the diagnosis, treatment and prevention of cell  
 CC proliferative, neurological and endocrine disorders, as well as in gene  
 CC therapy, antisense therapy to block transcription of the mRNA encoding  
 CC HPPN. HPPN polynucleotides may be used to generate hybridisation probes  
 CC useful for mapping the gene. HPPN may be used to treat or prevent a  
 CC neurological disorder comprising Alzheimer's disease, amyotrophic lateral  
 CC sclerosis, depression, diabetic neuropathy, epilepsy, Down's syndrome,  
 CC Huntington's disease and Parkinson's disease, an endocrine disorder  
 CC comprising Cushing's disease, diabetes insipidus, diabetes mellitus,  
 CC hyper- and hypoglycaemia and goitre. Agonists of HPPN may also be used to  
 CC treat or prevent the above mentioned disorders, whereas an antagonist  
 CC of HPPN may be used to treat or prevent a cell proliferative disorder  
 CC comprising arteriosclerosis, atherosclerosis, hepatitis and cancer  
 CC including leukaemia. Antibodies specific for HPPN may be used for the  
 CC diagnosis of the disorders or in an assay to monitor patients being  
 CC treated with HPPN or agonists, antagonists, and inhibitors of HPPN.  
 CC The present sequence represents HPPN and is encoded by a cDNA derived  
 CC from overlapping Incyte clones 2923961, 1576389 and 1309070.  
 XX  
 SQ Sequence 170 AA;  
 Query Match 100.0%; Score 890; DB 22; Length 170;  
 Best Local Similarity 100.0%; Pred. No. 1e-86;  
 Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MMAGMKIQLVCMLLAFSSWSLCSSEEMKALEADFLTNMHTSKISKAHVPSWKMTLLN 60  
 Db 1 MMAGMKIQLVCMLLAFSSWSLCSSEEMKALEADFLTNMHTSKISKAHVPSWKMTLLN 60  
 QY 61 VCSLVNNLNSPAEETGEVHEELVARRKLPALDGFSLSEAMLTIIYQLHKICHSAFOHWE 120  
 Db 61 VCSLVNNLNSPAEETGEVHEELVARRKLPALDGFSLSEAMLTIIYQLHKICHSAFOHWE 120  
 QY 121 LIQEDILTGDNDKNGKEEVIRKIPYILKQLYENKPRRPYILKRDSYYY 170  
 Db 121 LIQEDILTGDNDKNGKEEVIRKIPYILKQLYENKPRRPYILKRDSYYY 170  
 RESULT 2  
 ABUS6410  
 ID ABUS6410 standard; Protein; 170 AA.  
 XX  
 AC ABUS6410;  
 XX  
 DT 02-APR-2003 (first entry)  
 XX

DE Lung cancer-associated polypeptide #3.  
 XX  
 KW Lung cancer-associated polypeptide; cytostatic; emphysema;  
 KW antiinflammatory; antiasthmatic; non-small cell lung cancer; atelectasis;  
 KW small cell lung cancer; benign lesion; precancerous lesion; bronchitis;  
 KW chronic obstructive pulmonary disease; hypersensitivity pneumonitis;  
 KW interstitial pulmonary fibrosis; fibrosis; asthma; bronchiectasis.  
 XX  
 OS Unidentified.  
 XX  
 PN WO200286443-A2.  
 XX  
 PD 31-OCT-2002.  
 XX  
 PF 18-APR-2002; 2002WO-US12476.  
 XX  
 PR 18-APR-2001; 2001US-284770P.  
 PR 10-MAY-2001; 2001US-290492P.  
 PR 09-NOV-2001; 2001US-339245P.  
 PR 13-NOV-2001; 2001US-350666P.  
 PR 29-NOV-2001; 2001US-334370P.  
 PR 12-APR-2002; 2002US-372248P.  
 XX  
 PA (EOSB-) EOS BIOTECHNOLOGY INC.  
 XX  
 PI Aziz N, Murray R;  
 XX WPI; 2003-093161/08.  
 DR N-PSDB; ABX76126.  
 XX  
 CC Detecting a lung cancer-associated transcript in a cell from a patient  
 CC for treating lung cancer, by contacting a biological sample from the  
 CC patient with a polynucleotide that exhibits increased or decreased  
 CC expression in lung cancer -  
 XX  
 CC Claim 27; Page 190; 453pp; English.  
 CC  
 CC The invention relates to a method for detecting a lung cancer-associated  
 CC transcript in a cell from a patient, comprising contacting a biological  
 CC sample from the patient with a polynucleotide that selectively hybridises  
 CC to a sequence that is at least 80 % identical to a gene that exhibits  
 CC increased or decreased expression in lung cancer samples. Lung  
 CC cancer-associated polynucleotides and polypeptides are used for  
 CC identifying a compound that modulates a lung cancer-associated  
 CC polypeptide, for inhibiting proliferation of a lung cancer-associated  
 CC cell to treat lung cancer in a patient and for treating a mammal having  
 CC lung cancer by administering a modulatory compound identified. The  
 CC methods are useful for treating lung cancer, such as small cell lung  
 CC cancer, non-small cell lung cancer or other benign or precancerous  
 CC lesions, e.g. atelectasis, emphysema, bronchitis, chronic obstructive  
 CC pulmonary disease, fibrosis, hypersensitivity pneumonitis, interstitial  
 CC pulmonary fibrosis, asthma and bronchiectasis. The genes, polynucleotides  
 CC and polypeptides are useful for diagnostic purposes and as targets for  
 CC screening for therapeutic compounds that modulate lung cancer, such as  
 CC antibodies. Sequences ABUS6408-ABUS6745 represent lung cancer-associated  
 CC polypeptides of the invention.  
 XX  
 SQ Sequence 170 AA;  
 Query Match 100.0%; Score 890; DB 24; Length 170;  
 Best Local Similarity 100.0%; Pred. No. 1e-86;  
 Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MMAGMKIQLVCMLLAFSSWSLCSSEEMKALEADFLTNMHTSKISKAHVPSWKMTLLN 60  
 Db 1 MMAGMKIQLVCMLLAFSSWSLCSSEEMKALEADFLTNMHTSKISKAHVPSWKMTLLN 60  
 QY 61 VCSLVNNLNSPAEETGEVHEELVARRKLPALDGFSLSEAMLTIIYQLHKICHSAFOHWE 120  
 Db 61 VCSLVNNLNSPAEETGEVHEELVARRKLPALDGFSLSEAMLTIIYQLHKICHSAFOHWE 120  
 QY 121 LIQEDILTGDNDKNGKEEVIRKIPYILKQLYENKPRRPYILKRDSYYY 170





Qy 61 VCSSLNNSPAEETGEVHEELVARRKLPALDGFSLKAMLTIIYQLHKICHSAFOHWE 120  
 |||||  
 Db 61 VCSSLNNSPAEETGEVHEELVARRKLPALDGFSLKAMLTIIYQLHKICHSAFOHWE 120  
 |||||  
 Qy 121 LIQEDILDT 129  
 |||||  
 Db 121 LIQEDILDT 129

RESULT 5  
 AAB29865  
 ID AAB29865 standard; Protein; 279 AA.  
 XX  
 AC AAB29865;  
 XX  
 DT 09-FEB-2001 (first entry)  
 XX  
 DE Human secreted protein BLAST search protein SEQ ID NO: 123.

XX Cytostatic; immunosuppressive; neutropic; neuroprotective; antiviral;  
 KW antiallergic; hepatotropic; antidiabetic; antiinflammatory; antiulcer;  
 KW vulnary; anticonvulsant; antibacterial; antifungal; antiparasitic;  
 KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;  
 KW neurological disease; infection; human; secreted protein.

XX Homo sapiens.  
 OS  
 XX WO200061779-A1.  
 PN  
 XX 19-OCT-2000.  
 PD  
 XX 06-APR-2000; 2000WO-US09068.  
 PF  
 XX 09-APR-1999; 99US-0128699.  
 PR  
 XX 20-JAN-2000; 2000US-0177050.  
 PR  
 XX (HUMA-) HUMAN GENOME SCI INC.  
 PA

PI Rosen CA, Ruben SM, Komatsoulis G;  
 XX  
 XX WPI; 2000-647424/62.  
 DR

XX Isolated nucleic acid molecule encoding a human secreted protein is  
 PT used in preventing, treating or ameliorating a medical condition -

XX Disclosure; Page 457-458; 495pp; English.  
 XX  
 CC The invention relates to the isolation of genes AAC63410-C63458 encoding  
 CC the human secreted proteins AAB29802-B29850. This sequence represents a  
 CC peptide fragment homologous to the protein encoded by the gene given  
 CC in the descriptor line. The sequence is a search result from a BLASTX  
 CC homology search. The genes and proteins are useful for preventing,  
 CC ameliorating or treating medical conditions, e.g. by protein or gene  
 CC therapy. The genes are isolated from a range of human tissues disclosed  
 CC in the specification. The nucleic acids, proteins, antibodies and  
 CC (ant)agonists are useful in the diagnosis, treatment and prevention of:  
 CC (a) cancer, e.g. breast and ovarian cancer, and other cancers of the  
 CC adrenal gland, bone, bone marrow, breast, gastrointestinal tract, liver,  
 CC lung, or urogenital; (b) immune disorders e.g. Addison's disease,  
 CC allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis,  
 CC diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid  
 CC arthritis and ulcerative colitis; (c) cardiovascular disorders such  
 CC as myocardial ischaemia; (d) wound healing; (e) neurological diseases  
 CC e.g. cerebral anoxia and epilepsy; and (f) infectious diseases such  
 CC as viral, bacterial, fungal and parasitic infections.

XX Sequence 279 AA;  
 SQ  
 Query Match 10.1%; Score 90; DB 21; Length 279;  
 Best Local Similarity 22.2%; Pred. NO. 0.42;  
 Matches 24; Conservative 28; Mismatches 44; Indels 12; Gaps 3;

Qy 26 SEEMKALEADFLTNMHTSKISKARV-----PSWKMTLLNVCSSLNNSPAEETGEVH 79  
 |||||  
 Db 165 SEGKSNMPADLVLPQPSQVNSLEVITPETAETWRCVLTACKSLSHSI---GQELQVW 221  
 |||||  
 Qy 80 EEEIVARRKLPALDGFSLKAMLTIIYQLHKICHSR---AFQHWELIQE 124  
 |||||  
 Db 222 ENDLVAARQTRDSIGSQDLSRLLTWARMMSVSYGETTILSLHWMVLE 269

RESULT 6  
 AAY19902  
 ID AAY19902 standard; Protein; 322 AA.  
 XX  
 AC AAY19902;  
 XX  
 DT 19-JUL-1999 (first entry)  
 XX  
 DE B. burgdorferi antigenic protein, f210.aa.

XX Antigenic protein; vaccine; Lyme disease; infection; detection.  
 KW  
 XX Borrelia burgdorferi.

OS  
 XX WO9859071-A1.  
 PN  
 XX 30-DEC-1998.  
 PD

XX 18-JUN-1998; 98WO-US12718.  
 PF  
 XX 03-SEP-1997; 97US-0057483.  
 PR  
 XX 20-JUN-1997; 97US-0050359.  
 PR  
 XX 22-JUL-1997; 97US-0053344.  
 PR  
 XX 22-JUL-1997; 97US-0053377.

XX (HUMA-) HUMAN GENOME SCI INC.  
 PA  
 PA (MEDI-) MEDIMMUNE INC.

XX Choi GH, Erwin AL, Hanson MS, Lathigra R;  
 PI  
 XX WPI; 1999-189980/16.  
 DR  
 XX N-PSDB; AAX61599.

XX New isolated Borrelia burgdorferi nucleic acids - used to develop  
 PT products for the diagnosis, prevention and treatment of diseases  
 PT caused by Borrelia, particularly Lyme disease

XX Claim 12; Page 112; 275pp; English.

XX This sequence represents a Borrelia burgdorferi (Bb) protein of the  
 CC invention, which is suitable for use in a vaccine. The Bb polypeptides  
 CC can be used in vaccines for eliciting protective antibodies to members of  
 CC the Borrelia genus, particularly for the use against Lyme disease in  
 CC humans and animals. They can be used for preventing or attenuating an  
 CC infection caused by a member of the Borrelia genus. The products can also  
 CC be used for detection of members of the Borrelia genus.

XX Sequence 322 AA;

Query Match 9.8%; Score 87.5; DB 20; Length 322;  
 Best Local Similarity 25.9%; Pred. No. 0.95;  
 Matches 44; Conservative 32; Mismatches 55; Indels 39; Gaps 9;

Qy 5 MKIOLVCMLLAPF-----SWSLCSDBEEEMK-----ALEADFLTNMHTSKISKA 49  
 |||||  
 Db 1 MKIQLIMLLALLDFPLNARLLDISIKRADEIKYSSVNLILEKEYTTFPTSEIK- 59  
 |||||

Qy 50 HVPSSKMTLLNVCSSLVNN-----LNSPAETGE-VHEEELVARRKLPALDGFSLKAML 102  
 |||||

Db 60 --NIYKUTEHFVKSIMLNKNTYSLNSNYKANKYLQSELIDKKFLKYK--FKIKIN 115  
 |||||  
 Qy 103 TIYQLHKICHRS-RAFOHWELIQEDILDTGNDKNGKE--EVIKRKIPYILK 149  
 |||||  
 Db 116 GIFKSHSLIYTKGFGYKLELYIE-----NNAEPLKIFNLNITYFLK 156

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 29, 2003, 17:16:39 ; Search time 21 Seconds  
(without alignments)  
342.516 Million cell updates/sec

Title: US-09-927-565A-1

Perfect score: 890

Sequence: 1 MMAGKTKQLVCMLLAFSSW.....QLYENKPRPYLKRDYYY 170

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents\_AA.\*  
1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pep.\*  
2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pep.\*  
3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pep.\*  
4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pep.\*  
5: /cgn2\_6/ptodata/1/iaa/PCTUS\_COMB.pep.\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	890	100.0	170	3	US-09-002-114-1
2	804	90.3	169	3	US-09-002-114-3
3	711	79.9	169	3	US-09-002-114-4
4	80.5	9.0	550	3	US-09-166-460-13
5	80.5	9.0	550	3	US-09-361-718-13
6	80	9.0	424	4	US-09-198-452A-45
7	79.5	8.9	652	4	US-08-559-896B-2
8	78.5	8.8	550	3	US-09-166-460-2
9	78.5	8.8	550	3	US-09-166-460-5
10	78.5	8.8	550	3	US-09-166-460-7
11	78.5	8.8	550	3	US-09-166-460-9
12	78.5	8.8	550	3	US-09-166-460-11
13	78.5	8.8	550	3	US-09-166-460-15
14	78.5	8.8	550	3	US-09-166-460-17
15	78.5	8.8	550	3	US-09-166-460-19
16	78.5	8.8	550	3	US-09-166-460-20
17	78.5	8.8	550	3	US-09-166-460-21
18	78.5	8.8	550	3	US-09-166-460-23
19	78.5	8.8	550	3	US-09-166-460-24
20	78.5	8.8	550	3	US-09-166-460-25
21	78.5	8.8	550	3	US-09-166-460-28
22	78.5	8.8	550	3	US-09-361-718-2
23	78.5	8.8	550	3	US-09-361-718-5
24	78.5	8.8	550	3	US-09-361-718-9
25	78.5	8.8	550	3	US-09-361-718-11
26	78.5	8.8	550	3	US-09-361-718-15
27	78.5	8.8	550	3	US-09-361-718-15

28 78.5 8.8 550 3 US-09-361-718-17 Sequence 17, Appl  
29 78.5 8.8 550 3 US-09-361-718-19 Sequence 19, Appl  
30 78.5 8.8 550 3 US-09-361-718-20 Sequence 20, Appl  
31 78.5 8.8 550 3 US-09-361-718-21 Sequence 21, Appl  
32 78.5 8.8 550 3 US-09-361-718-23 Sequence 23, Appl  
33 78.5 8.8 550 3 US-09-361-718-24 Sequence 24, Appl  
34 78.5 8.8 550 3 US-09-361-718-25 Sequence 25, Appl  
35 78.5 8.8 550 3 US-09-361-718-28 Sequence 28, Appl  
36 78.5 8.8 550 4 US-09-398-395A-28 Sequence 28, Appl  
37 78.5 8.8 550 4 US-09-887-586A-28 Sequence 28, Appl  
38 78.5 8.8 550 4 US-09-895-752-28 Sequence 28, Appl  
39 78.5 8.8 550 4 US-09-903-012B-28 Sequence 28, Appl  
40 76.5 8.6 550 3 US-09-166-460-27 Sequence 27, Appl  
41 76.5 8.6 550 3 US-09-361-718-27 Sequence 27, Appl  
42 75 8.4 233 1 US-08-081-448-6 Sequence 6, Appl  
43 75 8.4 233 2 US-08-607-269-24 Sequence 24, Appl  
44 75 8.4 233 2 US-08-470-670A-7 Sequence 7, Appl  
45 75 8.4 233 3 US-08-167-921-2 Sequence 2, Appl

#### ALIGNMENTS

RESULT 1  
US-09-002-114-1  
; Sequence 1, Application US/09002114  
; Patent No. 6274720  
; GENERAL INFORMATION:  
; APPLICANT: Lal, Preeti  
; APPLICANT: Shah, Purvi  
; APPLICANT: Corley, Neil C.  
; TITLE OF INVENTION: HUMAN PREPRONEUROTENSIN/NEUROMEDIN N  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Incyte Pharmaceuticals, Inc.  
; STREET: 3174 Porter Drive  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA: US/09/002,114  
; APPLICATION NUMBER: US/09/002,114  
; FILING DATE: Herewith  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Billings, Lucy J.  
; REGISTRATION NUMBER: 36,749  
; REFERENCE/DOCKET NUMBER: PF-0450 US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 650-855-0555  
; TELEFAX: 650-845-4166  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 170 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; IMMEDIATE SOURCE:  
; LIBRARY: PITUNOT03  
; CLONE: 1760566  
; US-09-002-114-1  
Query Match 100.0%; Score 890; DB 3; Length 170;  
Best Local Similarity 100.0%; Pred. No. 1e-95;

Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MMAGKIQVCMVLLAFSSWSLCSDEEMKALEADFLTNMHTSKISKAHVPSWKMTLLN 60  
Db 1 MMAGKIQVCMVLLAFSSWSLCSDEEMKALEADFLTNMHTSKISKAHVPSWKMTLLN 60  
Qy 61 VCSLVNLSNLSPAEETGEVHEELVARRKLPALDGFSLAEMLTIIYQLHKICHSAFOHWE 120  
Db 61 VCSLVNLSNLSPAEETGEVHEELVARRKLPALDGFSLAEMLTIIYQLHKICHSAFOHWE 120  
Qy 121 LIQEDILTDGNDKNGKEVIRKIPYILKQLYENKPRRPYILKRDSSYY 170  
Db 121 LIQEDILTDGNDKNGKEVIRKIPYILKQLYENKPRRPYILKRDSSYY 170

## RESULT 2

US-09-002-114-3  
; Sequence 3, Application US/09002114  
; Patent No. 6274720  
; GENERAL INFORMATION:  
; APPLICANT: Lal, Preeti  
; APPLICANT: Shah, Purvi  
; APPLICANT: Corley, Neil C.  
; TITLE OF INVENTION: HUMAN PREPRONEUROTENSIN/NEUROMEDIN N  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Incyte Pharmaceuticals, Inc.  
; STREET: 3174 Porter Drive  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94304

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/002,114  
FILING DATE: Herewith

## CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Billings, Lucy J.

REGISTRATION NUMBER: 36,749

REFERENCE/DOCKET NUMBER: PF-0450 US

TELECOMMUNICATION INFORMATION:

TELEPHONE: 650-855-0555

TELEFAX: 650-845-4166

TELEX:

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 169 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

IMMEDIATE SOURCE:

LIBRARY: GenBank

CLONE: 163424

US-09-002-114-3

Query Match 90.3%; Score 804; DB 3; Length 169;

Best Local Similarity 90.5%; Pred. NO. 1e-85;

Matches 153; Conservative 4; Mismatches 12; Indels 0; Gaps 0;

Qy 2 MAGMKIQVCMVLLAFSSWSLCSDEEMKALEADFLTNMHTSKISKAHVPSWKMTLLN 61  
Db 1 MAGMKIQVCMVLLAFSSWSLCSDEEMKALEADFLTNMHTSKISKASVPSWKMTLLN 60  
Qy 62 CSLVNNLSNLSPAEETGEVHEELVARRKLPALDGFSLAEMLTIIYQLHKICHSAFOHWE 121

Db 61 CSLINNLSQAETGEFHEBELITRRKFPALDGFSLAEMLTIIYQLKICHSAFOHWE 120  
Qy 122 IOEDILTDGNDKNGKEVIRKIPYILKQLYENKPRRPYILKRDSSYY 170  
Db 121 IOEDILTDGNDKNGKEVIRKIPYILKQLYENKPRRPYILKRDSSYY 169

## RESULT 3

US-09-002-114-4  
; Sequence 4, Application US/09002114  
; Patent No. 6274720  
; GENERAL INFORMATION:  
; APPLICANT: Lal, Preeti  
; APPLICANT: Shah, Purvi  
; APPLICANT: Corley, Neil C.  
; TITLE OF INVENTION: HUMAN PREPRONEUROTENSIN/NEUROMEDIN N  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Incyte Pharmaceuticals, Inc.  
; STREET: 3174 Porter Drive  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94304

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/002,114  
FILING DATE: Herewith

## CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Billings, Lucy J.

REGISTRATION NUMBER: 36,749

REFERENCE/DOCKET NUMBER: PF-0450 US

TELECOMMUNICATION INFORMATION:

TELEPHONE: 650-855-0555

TELEFAX: 650-845-4166

TELEX:

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 169 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

IMMEDIATE SOURCE:

LIBRARY: GenBank

CLONE: 92546

US-09-002-114-4

Query Match 79.9%; Score 711; DB 3; Length 169;

Best Local Similarity 78.1%; Pred. NO. 6.9e-75;

Matches 132; Conservative 17; Mismatches 20; Indels 0; Gaps 0;

Qy 2 MAGMKIQVCMVLLAFSSWSLCSDEEMKALEADFLTNMHTSKISKAHVPSWKMTLLN 61  
Db 1 MIGNVQLVCLTLAFSSWSLCSDEEDVRALEADLTNNHASKVSKGSPSPWKMTLLN 60  
Qy 62 CSLVNNLSNLSPAEETGEVHEELVARRKLPALDGFSLAEMLTIIYQLHKICHSAFOHWE 121  
Db 61 CSLINNLSAAEEAGEMRDDDLVAKRKLPLVLDLDFSLAELLTVFQLKICRSRAFOHWEI 120

Qy 122 IOEDILTDGNDKNGKEVIRKIPYILKQLYENKPRRPYILKRDSSYY 170  
Db 121 IOEDILTDGNDKNGKEVIRKIPYILKQLYENKPRRPYILKRDSSYY 169

## RESULT 4

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

# OM protein - protein search, using sw model

Run on: December 29, 2003, 17:17:34 ; Search time 32 Seconds  
(without alignments)  
1057.280 Million cell updates/sec

Title: US-09-927-565A-1

Perfect score: 890

Sequence: 1 MMAGMKIQLVCMLLAFSSW.....QLYENKPRRPYILKRDSSYY 170

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 724715 seqs, 199017464 residues

Total number of hits satisfying chosen parameters: 724715

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

## Database :

Published Applications AA:\*  
1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*  
17: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*  
18: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	890	100.0	170	12	US-10-241-220-111
2	890	100.0	170	12	US-10-295-027-316
3	82	9.2	400	12	US-10-369-493-18274
4	80.5	9.0	403	12	US-10-369-493-20145
5	80.5	9.0	550	12	US-09-860-282A-13
6	79.5	8.9	652	9	US-09-351-794A-2
7	78.5	8.8	550	9	US-09-887-586A-28
8	78.5	8.8	550	9	US-09-903-012-28
9	78.5	8.8	550	11	US-09-900-797-28
10	78.5	8.8	550	12	US-09-860-282A-2
11	78.5	8.8	550	12	US-09-860-282A-5
12	78.5	8.8	550	12	US-09-860-282A-7
13	78.5	8.8	550	12	US-09-860-282A-9
14	78.5	8.8	550	12	US-09-860-282A-11
15	78.5	8.8	550	12	US-09-860-282A-15

16	78.5	8.8	550	12	US-09-860-282A-17	Sequence 17, Appl
17	78.5	8.8	550	12	US-09-860-282A-19	Sequence 19, Appl
18	78.5	8.8	550	12	US-09-860-282A-20	Sequence 20, Appl
19	78.5	8.8	550	12	US-09-860-282A-21	Sequence 21, Appl
20	78.5	8.8	550	12	US-09-860-282A-23	Sequence 23, Appl
21	78.5	8.8	550	12	US-09-860-282A-24	Sequence 24, Appl
22	78.5	8.8	550	12	US-09-860-282A-25	Sequence 25, Appl
23	78.5	8.8	550	12	US-09-860-282A-28	Sequence 28, Appl
24	76.5	8.6	216	12	US-10-056-479A-1	Sequence 1, Appl
25	76.5	8.6	550	12	US-09-860-282A-27	Sequence 27, Appl
26	75.5	8.5	491	15	US-10-128-714-3267	Sequence 3267, Ap
27	75.5	8.5	554	15	US-10-128-714-8267	Sequence 8267, Ap
28	75	8.4	212	12	US-10-169-223-14	Sequence 14, Appl
29	75	8.4	233	9	US-09-734-846-2	Sequence 2, Appl
30	75	8.4	233	10	US-09-952-278-6	Sequence 6, Appl
31	75	8.4	233	12	US-10-169-223-10	Sequence 10, Appl
32	75	8.4	233	12	US-10-302-262-2	Sequence 2, Appl
33	75	8.4	233	12	US-10-116-275-171	Sequence 171, App
34	75	8.4	233	15	US-10-072-830-4	Sequence 4, Appl
35	75	8.4	233	14	US-10-029-180-96	Sequence 96, Appl
36	75	8.4	1097	12	US-10-032-585-7036	Sequence 7036, Ap
37	74.5	8.4	550	12	US-09-860-282A-22	Sequence 22, Appl
38	74	8.3	405	15	US-10-106-698-6333	Sequence 6333, Ap
39	74	8.3	662	9	US-09-799-777-53	Sequence 53, Appl
40	74	8.3	671	12	US-10-311-626-5	Sequence 5, Appl
41	73.5	8.3	1038	12	US-10-032-585-7776	Sequence 7776, Ap
42	73	8.2	552	10	US-09-961-721-2	Sequence 2, Appl
43	73	8.2	552	12	US-10-170-789-43	Sequence 43, Appl
44	72.5	8.1	235	12	US-10-369-493-11220	Sequence 11220, A
45	72.5	8.1	461	15	US-10-234-432-24	Sequence 24, Appl

## ALIGNMENTS

RESULT 1  
US-10-241-220-111  
; Sequence 111, Application US/10241220  
; Publication No. US20030148408A1  
; GENERAL INFORMATION:  
; APPLICANT: Frantz, Gretchen  
; APPLICANT: Hillan, Kenneth J.  
; APPLICANT: Phillips, Heidi  
; APPLICANT: Polakis, Paul  
; APPLICANT: Spencer, Susan  
; APPLICANT: Williams, P. Mickey  
; APPLICANT: Wu, Thomas  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND  
; FILE OF INVENTION: TREATMENT OF TUMOR  
; FILE REFERENCE: P5010R1-US  
; CURRENT APPLICATION NUMBER: US/10/241,220  
; CURRENT FILING DATE: 2002-12-13  
; NUMBER OF SEQ ID NOS: 120  
; SEQ ID NO 111  
; LENGTH: 170  
; TYPE: PRT  
; ORGANISM: Homo Sapien  
US-10-241-220-111

Query Match 100.0%; Score 890; DB 12; Length 170;  
Best Local Similarity 100.0%; Pred. No. 4e-90;  
Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MMAGMKIQLVCMLLAFSSW	SLCSDSEEMKALBDFLTNNHTSKISKAHVPSWMTLLN	60
Db	1	MMAGMKIQLVCMLLAFSSW	SLCSDSEEMKALBDFLTNNHTSKISKAHVPSWMTLLN	60
Qy	61	VCSLVNNLNSPAETGVEVHEELVARBKLP	TALDGFSLKMLTIYQLHKICHSAFAQHW	120
Db	61	VCSLVNNLNSPAETGVEVHEELVARBKLP	TALDGFSLKMLTIYQLHKICHSAFAQHW	120
Qy	121	LIQEDILTGDNGKNGKEVIKPKIPYILK	RLQYENKPRRPYILKRDSSYY	170

Db 121 LIQEDILDTGNDKNGKEVIRKIPYILKQYENKPRPYILKEDSYYY 170

## RESULT 2

US-10-295-027-316

; Sequence 316, Application US/10295027

; Publication No. US20030232350A1

; GENERAL INFORMATION:

; APPLICANT: Afar, Daniel

; APPLICANT: Aziz, Natasha

; APPLICANT: Ginsberg, Wendy M.

; APPLICANT: Gish, Kurt C.

; APPLICANT: Glynn, Richard

; APPLICANT: Hevezi, Peter A.

; APPLICANT: Mack, David H.

; APPLICANT: Murray, Richard

; APPLICANT: Watson, Susan R.

; APPLICANT: Eos Biotechnology, Inc.

; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and

; TITLE OF INVENTION: Methods of Screening for Modulators of Cancer

; FILE REFERENCE: 018501-012500US

; CURRENT APPLICATION NUMBER: US/10/295,027

; CURRENT FILING DATE: 2002-11-13

; PRIOR APPLICATION NUMBER: US 09/663,733

; PRIOR FILING DATE: 2000-09-15

; PRIOR FILING DATE: 2001-11-13

; PRIOR FILING DATE: 2001-11-13

; PRIOR FILING DATE: 2001-11-15

; PRIOR FILING DATE: 2001-11-15

; PRIOR FILING DATE: 2001-11-21

; PRIOR FILING DATE: 2001-11-21

; PRIOR FILING DATE: 2001-11-21

; PRIOR FILING DATE: 2001-11-21

; PRIOR FILING DATE: 2001-11-21

; PRIOR FILING DATE: 2001-11-21

; PRIOR FILING DATE: 2001-11-21

; PRIOR FILING DATE: 2001-11-21

; PRIOR FILING DATE: 2001-11-21

; PRIOR FILING DATE: 2001-11-21

; PRIOR FILING DATE: 2001-11-21

; PRIOR FILING DATE: 2001-11-21

; PRIOR FILING DATE: 2001-11-21

; PRIOR FILING DATE: 2001-11-21

; PRIOR FILING DATE: 2001-11-21

; PRIOR FILING DATE: 2001-11-21

; PRIOR FILING DATE: 2001-11-21

; PRIOR FILING DATE: 2001-11-21

; PRIOR FILING DATE: 2001-11-21

; PRIOR FILING DATE: 2001-11-21

; PRIOR FILING DATE: 2001-11-21

; PRIOR FILING DATE: 2001-11-21

; PRIOR FILING DATE: 2001-11-21

; PRIOR FILING DATE: 2001-11-21

; PRIOR FILING DATE: 2001-11-21

; PRIOR FILING DATE: 2001-11-21

; PRIOR FILING DATE: 2001-11-21

; PRIOR FILING DATE: 2001-11-21

; PRIOR FILING DATE: 2001-11-21

; PRIOR FILING DATE: 2001-11-21

; PRIOR FILING DATE: 2001-11-21

; PRIOR FILING DATE: 2001-11-21

; PRIOR FILING DATE: 2001-11-21

; PRIOR FILING DATE: 2001-11-21

; PRIOR FILING DATE: 2001-11-21

; PRIOR FILING DATE: 2001-11-21

; PRIOR FILING DATE: 2001-11-21

; PRIOR FILING DATE: 2001-11-21

; PRIOR FILING DATE: 2001-11-21

; PRIOR FILING DATE: 2001-11-21

; PRIOR FILING DATE: 2001-11-21

; PRIOR FILING DATE: 2001-11-21

; PRIOR FILING DATE: 2001-11-21

; PRIOR FILING DATE: 2001-11-21

; PRIOR FILING DATE: 2001-11-21

; PRIOR FILING DATE: 2001-11-21

; PRIOR FILING DATE: 2001-11-21

; PRIOR FILING DATE: 2001-11-21

; PRIOR FILING DATE: 2001-11-21

; PRIOR FILING DATE: 2001-11-21

## RESULT 3

US-10-369-493-18274

; Sequence 18274, Application US/10369493

; Publication No. US20030233675A1

; GENERAL INFORMATION:

; APPLICANT: Cao, Yongwei

; APPLICANT: Hinkle, Gregory J.

; APPLICANT: Slater, Steven C.

; APPLICANT: Goldman, Barry S.

; APPLICANT: Chen, Xianfeng

; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF

; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES

; FILE REFERENCE: 38-10(52052)B

; CURRENT APPLICATION NUMBER: US/10/369,493

; CURRENT FILING DATE: 2003-02-28

; PRIOR APPLICATION NUMBER: US 60/360,039

; PRIOR FILING DATE: 2002-02-21

; NUMBER OF SEQ ID NOS: 47374

; SEQ ID NO 18274

; LENGTH: 400

; TYPE: PRT

; ORGANISM: Thermoplasma acidophilum

US-10-369-493-18274

Query Match 9.2%; Score 82; DB 12; Length 400;

Best Local Similarity 27.5%; Pred. No. 2;

Matches 33; Conservative 20; Mismatches 43; Indels 24; Gaps 6;

Qy 70 SPAETGEVHEELVARRKLPALDGFSLKAWLTIYQLHKICHSR-----AFQHWELIQ 123

Db 155 SPAPSAGKA--EAVHTAPQIPAKPAGRGSEILEMHGURRIIFDKMTKAKQIMPHFTVME 212

Qy 124 E-----DILDTGNDKNGKEEV---IKRKIPYILKQYEN---KPRRPYILKRDSSYY 169

Db 213 EVDVTSVMSILDSAKARNKVTVTGFLARIVPSILKQYPLNAYDETRVYILKK--YY 270

## RESULT 4

US-10-369-493-20145

; Sequence 20145, Application US/10369493

; Publication No. US20030233675A1

; GENERAL INFORMATION:

; APPLICANT: Cao, Yongwei

; APPLICANT: Hinkle, Gregory J.

; APPLICANT: Slater, Steven C.

; APPLICANT: Goldman, Barry S.

; APPLICANT: Chen, Xianfeng

; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF

; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES

; FILE REFERENCE: 38-10(52052)B

; CURRENT APPLICATION NUMBER: US/10/369,493

; CURRENT FILING DATE: 2003-02-28

; PRIOR APPLICATION NUMBER: US 60/360,039

; PRIOR FILING DATE: 2002-02-21

; NUMBER OF SEQ ID NOS: 47374

; SEQ ID NO 20145

; LENGTH: 403

; TYPE: PRT

; ORGANISM: No. US20030233675A1loc punctiforme

US-10-369-493-20145

Query Match 9.0%; Score 80.5; DB 12; Length 403;

Best Local Similarity 26.1%; Pred. No. 3;

Matches 36; Conservative 19; Mismatches 50; Indels 33; Gaps 7;

Qy 48 KAHVPSPKMTLLNVCSLVNNSLSPAEET--GEVHEELVAR---RKLPALDGFSLKAWL 102

Db 137 KVQRFGLDTIAQDIALQGIADLVARTFQGTQYIKSIAEFTKALEAELD--FTREAGF 195

Qy 103 TIYQLHKICHSRAFQ-----HWELIQDIL-----DTGNDKNGKEEVI 140

Db 196 TDQURRLSKSRWFDPTQIVVAEINWELTEKLLVMEWLDGVPFLSADLNNNNGKPAV 255

Qy 141 KRK-IPYILKR-----QLY 153

Db 256 ERKEITLLFRVFFQQLY 273

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

# OM protein - protein search, using sw model

Run on: December 29, 2003, 17:15:23 ; Search time 34 Seconds  
(without alignments)  
1290.263 Million cell updates/sec

Title: US-09-927-565A-1  
Perfect score: 890  
Sequence: 1 MMAGMKIQLVCMILLAFSSW.....OLYENKPRPYILKRDSYY 170

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

## Database :

SPTREMBL 23:\*

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phase:\*
- 10: sp\_plant:\*
- 11: sp\_rodent:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*
- 15: sp\_virus:\*
- 16: sp\_bacteriap:\*
- 17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Length	DB ID	Description
1	91.5	10.3	832 17 Q8PUW8	Q8PUW8 methanosarc
2	90	10.1	550 10 Q22880	Q22880 arabidopsis
3	87.5	9.8	171 16 O51616	O51616 borrelia bu
4	86.5	9.7	499 16 O67377	O67377 aquifex ao
5	86.5	9.7	605 12 Q9WT50	Q9WT50 human herpe
6	86.5	9.7	610 12 Q9Q349	Q9Q349 human herpe
7	84.5	9.5	651 16 Q8NV84	Q8NV84 staphylococ
8	84.5	9.5	979 5 Q95X88	Q95X88 caenorhabdi
9	84.5	9.5	1251 5 O16637	O16637 caenorhabdi
10	84	9.4	673 16 Q8EB90	Q8EB90 shewanella
11	82	9.2	400 17 Q9HTA5	Q9HTA5 thermoplas
12	82	9.2	805 16 Q8R8R1	Q8R8R1 thermoplas
13	81.5	9.2	652 5 Q9U0D4	Q9U0D4 plasmodium
14	81.5	9.2	1964 5 Q8IIX6	Q8IIX6 plasmodium
15	81	9.1	475 17 Q8U2Y3	Q8U2Y3 pyrococcus
16	81	9.1	510 10 Q9LJY7	Q9LJY7 arabidopsis

17	80	9.0	1229	16	Q8DTJ6	Q8dtj6 streptococ
18	79.5	8.9	479	12	Q8QQ67	Q8qq67 camelbox vi
19	79.5	8.9	479	12	Q8V2W4	Q8v2w4 camelbox vi
20	79.5	8.9	479	12	Q8V530	Q8v530 monkeypox v
21	79.5	8.9	479	12	Q85370	Q85370 variola vir
22	79.5	8.9	479	12	Q8JL94	Q8jl94 ectromelia
23	79.5	8.9	479	12	Q8JL94	Q8jl94 ectromelia
24	79.5	8.9	479	12	Q57184	Q57184 vaccinia vi
25	79.5	8.9	479	12	Q9JFD9	Q9jfd9 vaccinia vi
26	79.5	8.9	479	12	Q8QM29	Q8qm29 cowpox viru
27	79.5	8.9	479	12	Q8QX5	Q8qx5 staphylococ
28	79	8.9	229	10	Q9LSQ8	Q9lsq8 arabidopsis
29	79	8.9	741	11	Q8C0X0	Q8c0x0 mus musculu
30	78.5	8.8	312	13	Q8DGH8	Q8dgh8 gallus gall
31	78.5	8.8	479	12	Q9PXS1	Q9pxs1 variola vir
32	78.5	8.8	550	10	Q48935	Q48935 mentha pipe
33	78.5	8.8	913	11	Q9CYM1	Q9cym1 mus musculu
34	78	8.8	227	16	Q8FEL7	Q8fel7 corynebacte
35	78	8.8	421	5	Q8IH27	Q8ih27 plasmodium
36	78	8.8	537	6	Q29016	Q29016 sus scrofa
37	78	8.8	837	10	Q9LZC7	Q9lzc7 arabidopsis
38	78	8.8	896	10	Q9FEN1	Q9fen1 arabidopsis
39	78	8.8	908	10	Q8W4J9	Q8w4j9 arabidopsis
40	78	8.8	908	10	Q9ZSY3	Q9zsy3 arabidopsis
41	78	8.8	1849	5	Q8I3L3	Q8i3l3 plasmodium
42	77.5	8.7	270	2	Q9EU59	Q9eu59 chlamydia t
43	77.5	8.7	654	4	Q8WX71	Q8wx71 homo sapien
44	77.5	8.7	1793	4	Q8TEP3	Q8tep3 homo sapien
45	77	8.7	346	2	Q60044	Q60044 thermotoga

## ALIGNMENTS

## RESULT 1

ID	Q8PUW8	PRELIMINARY;	PRT;	832 AA.
AC	Q8PUW8;			
DC	01-OCT-2002 (TREMBLrel. 22, Created)			
DT	01-OCT-2002 (TREMBLrel. 22, Last sequence update)			
DE	01-MAR-2003 (TREMBLrel. 23, Last annotation update)			
GN	MM2210.			
OS	Methanosarcina mazei (Methanosarcina frisia).			
OC	Archaea; Euryarchaeota; Methanococci; Methanosarcinales;			
CC	Methanosarcinaceae; Methanosarcina.			
OX	NCBI_TaxID=2209;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=Goel / Go1 / ATCC BAA-199 / DSM 3647 / OCM 88;			
RX	MEDLINE=22120827; PubMed=12125824;			
RA	Deppeinmer U., Johann A., Hartsch T., Merkl R., Schmitz R.A.,			
RA	Martinez-Arias R., Henne A., Wieser A., Baeumer S., Jacobi C.,			
RA	Brueggemann H., Lienard T., Christmann A., Boemcke M., Steckel S.,			
RA	Bhattacharya A., Lykidis A., Overbeek R., Klenk H.-P., Gunsalus R.P.,			
RA	Fritz H.-J., Gottschalk G.			
RT	"The genome of Methanosarcina mazei: evidence for lateral gene			
RT	transfer between Bacteria and Archaea."			
RL	J. Mol. Microbiol. Biotechnol. 4:453-461 (2002).			
DR	EMBL; AE013461; AAM31906.1; -			
DR	InterPro; IPR003674; Oligo_trans_STT3.			
DR	Pfam; PF02516; STT3; 1.			
KW	Transferase; Complete proteome.			
SQ	SEQUENCE 832 AA; 92238 MW; DB862E2CBFAAFF34 CRC64;			

Query Match 10.3%; Score 91.5; DB 17; Length 832;

Best Local Similarity 27.4%; Pred. No. 2.8;

Matches 34; Conservative 24; Mismatches 47; Indels 19; Gaps 6;

Qy 22 LCSDBSEEMKALEADFLTNMHTSKISKAHVPSWQVTLNVCVSLVNNLNSPAETGEVHVE 81

Db 621 LLTDSBEAKTI-----VEKLKVYVMTDTLMAEGKPSITSLSL--AGKEIGYYEV 669

```

QY 82 ELVARRKLPALDGFSL--EAML--TIYQLHKICHSAFQHWELIQEDILDTGNDKNGKE 137
DB 670 ETV---KGDYGLRTVATPKQALLQVYKHLK-DGSLGHPRLVHSAVSTDDGNSKE 725
QY 138 EVIK 141
DB 726 NTVK 729

RESULT 2
O22880 PRELIMINARY; PRT; 550 AA.
AC O22880;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 15, Last annotation update)
DE AT2940550 protein.
GN AT2940550.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; Core eudicots; Rosidae;
OC eucosids II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=20083487; PubMed=10617197;
RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblyum T.V.,
RA Buell C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H., Moffat K.S.,
RA Cronin L.A., Shen M., VanAken S.E., Umayam L., Tallon L.J., Gill J.E.,
RA Adams M.D., Carrera A.J., Creasy T.H., Goodman H.M., Somerville C.R.,
RA Copenhaver G.P., Preuss D., Niernan W.C., White O., Eisen J.A.,
RA Salzberg S.L., Fraser C.M., Venter J.C.;
RT "Sequence and analysis of chromosome 2 of the plant Arabidopsis
thaliana.";
RL Nature 402:761-768(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RA Lin X.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC002336; AAB87582.1;
SQ SEQUENCE 550 AA; 61321 MW; 721153E2D5FB2126 CRC64;

Query Match 10.1%; Score 90; DB 10; Length 550;
Best Local Similarity 22.2%; Pred. No. 2.4;
Matches 24; Conservative 28; Mismatches 44; Indels 12; Gaps 3;

QY 26 SEEMKALEADPLTNMHTSKSKAHV-----PSWKMTLLNVCSLVNLSNPAETGEVH 79
DB 436 SEGKSNIPADLVLPQSPQNSLEVIPTETAETWRCYLATCKSLSHSI---GQELQGV 492
QY 80 BEELVARRKLPALDGFSL--EAML--TIYQLHKICHSAFQHWELIQE 124
DB 493 ENDLVARRKQLDSGLSQSLRLTLTWRMWSVSGYGTLLSLEHWQVLE 540

RESULT 3
O51616 PRELIMINARY; PRT; 171 AA.
AC O51616;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Hypothetical protein BB0673.
GN BB0673.
OS Borrelia burgdorferi (Lyme disease spirochete).
OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Borrelia.
OX NCBI_TaxID=139;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 35210 / B31;

```

```

RX MEDLINE=98065943; PubMed=9403685;
RA Fraser C.M., Castens S., Huang W.M., Sutton G.G., Clayton R.A.,
RA Lachigra B., White O., Ketchum K.A., Dodson R., Hickey E.K., Winn M.,
RA Dougherty B., Tomb J.-F., Fleischmann R.D., Richardson D.,
RA Peterson J., Kerlavage A.R., Quackenbush J., Salzberg S., Hanson M.,
RA van Vugt R., Palmer N., Adams M.D., Gocayne J.D., Weidman J.,
RA Utterback T., Matthey L., McDonald L., Artiach P., Bowman C.,
RA Garland S., Fujii C., Cotton M.D., Horst K., Roberts K., Hatch B.,
RA Smith H.O., Venter J.C.;
RT "Genomic sequence of a Lyme disease spirochete, Borrelia
burgdorferi.";
RL Nature 390:580-586(1997).
DR EMBL; AE001168; AAC67020.1;
DS TIGR; BB0673;
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 171 AA; 20537 MW; 8F7815D05D94E8F6 CRC64;

Query Match 9.8%; Score 87.5; DB 16; Length 171;
Best Local Similarity 25.9%; Pred. No. 1.1;
Matches 44; Conservative 32; Mismatches 55; Indels 39; Gaps 9;

QY 5 MKIQLVCMLLAFS-----SWSLCSDSSEEMK-----ALEADPLTNMHTSKSKA 49
DB 1 MKIQLIIMLLALLDFPLNARLLDISIEKRADEIKYSSYNLILEKEYYTNFTSEIK- 59
QY 50 HVPWKMTLLNVCSLVN-----LNSPAETGE-VHEEELVARRKLPALDGFSL 102
DB 60 --NTYKUTEHFVKSIMLNKNTYSLNSNYKANKYLIQSELDIKKFLKYKI--PKRNK 115
QY 103 TIYQLHKICHSAFQHWELIQEDILDTGNDKNGKE--EVIKRPIPYLK 149
DB 116 GIFKSHSLIYTKGFGYKLELYE-----NNAEPLKFNLTITFLK 156

RESULT 4
O67377 PRELIMINARY; PRT; 499 AA.
AC O67377;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Hypothetical protein AQ_1365.
GN AQ_1365.
OS Aquifex aeolicus.
OC Bacteria; Aquificae; Aquificales; Aquificaceae; Aquifex.
OX NCBI_TaxID=63363;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VF5;
RX MEDLINE=98196666; PubMed=9537320;
RA Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,
RA Graham D.E., Overbeek R., Sneed M.A., Keller M., Aufay M., Huber R.,
RA Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;
RT "The complete genome of the hyperthermophilic bacterium Aquifex
aeolicus.";
RL Nature 392:353-358(1998).
DR EMBL; AE000736; AAC07340.1;
DR InterPro; IPR000631; UPF0031.
DR Pfam; PF01256; UPF0031.1;
DR Pfam; PF03853; Yjef N; 1;
DR TIGRPFAMs; TIGR00196; Yjef cterm; 1;
DR TIGRPFAMs; TIGR00197; Yjef Nterm; 1;
DR PROSITE; PS01050; UPF0031.2; 1;
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 499 AA; 54371 MW; EDD6B5230AC50F3 CRC64;

Query Match 9.7%; Score 86.5; DB 16; Length 499;
Best Local Similarity 26.5%; Pred. No. 4.8;
Matches 35; Conservative 14; Mismatches 40; Indels 43; Gaps 5;

QY 52 PSWKMTLLNVCSLVNLSN-----AEETGEVHEEELVARRKLPALDGFSL 101

```



=> s (neurotensin? precursor) or proneurotensin? or preproneurotensin?  
L1 83 (NEUROTENSIN? PRECURSOR) OR PRONEUROTENSIN? OR PREPRONEUROTENSIN  
?

=> s 11 (12a) human  
L2 11 L1 (12A) HUMAN

=> duplicate remove  
ENTER L# LIST OR (END):12  
DUPLICATE PREFERENCE IS 'BIOSIS, USPATFULL, PCTFULL'  
KEEP DUPLICATES FROM MORE THAN ONE FILE? Y/(N):n  
PROCESSING COMPLETED FOR L2  
L3 10 DUPLICATE REMOVE L2 (1 DUPLICATE REMOVED)

=> d 1-10

L3 ANSWER 1 OF 10 USPATFULL on STN  
AN 2003:154406 USPATFULL  
TI Collections of transgenic animal lines (living library)  
IN Serafini, Tito Andrew, San Mateo, CA, UNITED STATES  
PI US 2003106074 A1 20030605  
AI US 2002-77025 A1 20020214 (10)  
RLI Continuation-in-part of Ser. No. US 2001-783487, filed on 14 Feb 2001,  
PENDING  
DT Utility  
FS APPLICATION  
LN.CNT 5667  
INCL INCLM: 800/008.000  
INCLS: 800/014.000  
NCL NCLM: 800/008.000  
NCLS: 800/014.000  
IC [7]  
ICM: A01K067-033  
ICS: A01K067-027  
CAS INDEXING IS AVAILABLE FOR THIS PATENT.

L3 ANSWER 2 OF 10 USPATFULL on STN  
AN 2003:72979 USPATFULL  
TI Collections of transgenic animal lines (living library)  
IN Serafini, Tito Andrew, San Mateo, CA, UNITED STATES  
PI US 2003051266 A1 20030313  
AI US 2001-783487 A1 20010214 (9)  
DT Utility  
FS APPLICATION  
LN.CNT 4818  
INCL INCLM: 800/018.000  
INCLS: 435/007.100  
NCL NCLM: 800/018.000  
NCLS: 435/007.100  
IC [7]  
ICM: A01K067-027  
ICS: G01N033-53  
CAS INDEXING IS AVAILABLE FOR THIS PATENT.

L3 ANSWER 3 OF 10 PCTFULL COPYRIGHT 2004 Univentio on STN  
AN 2003038049 PCTFULL ED 20030515 EW 200319  
TIEN METHOD FOR ISOLATING CELL-TYPE SPECIFIC MRNAS  
TIFR METHODE PERMETTANT D'ISOLER DES ARNM SPECIFIQUES D'UN TYPE DE CELLULE  
IN HEINTZ, Nathaniel, 441 Fowler Avenue, Pelham Manor, NY 10803, US [US,  
US];  
SERAFINI, Tito, A., 1835 Mulberry Drive, San Mateo, CA 94403, US [US,  
US];  
SHYJAN, Andrew, W., 177 Barford Avenue, San Carlos, CA 94070, US [US,  
US]  
PA RENOVIS, INC., 270 Littlefield Avenue, South San Francisco, CA 94080, US  
[US, US], for all designates States except US;  
HEINTZ, Nathaniel, 441 Fowler Avenue, Pelham Manor, NY 10803, US [US,  
US], for US only;

**THIS PAGE BLANK (USPTO)**

SERAFINI, Tito, A., 1835 Mulberry Drive, San Mateo, CA 94403, US [US,  
US], for US only;  
SHYJAN, Andrew, W., 177 Barford Avenue, San Carlos, CA 94070, US [US,  
US], for US only  
AG ANTLER, Adriane, M., Pennie & Edmonds LLP, 1155 Avenue of the Americas,  
New York, NY 10036, US  
LAF English  
LA English  
DT Patent  
PI WO 2003038049 A2 20030508  
DS W: AE AG AL AM AT AU AZ BA BB BG BR BY BZ CA CH CN CO CR CU CZ  
DE DK DM DZ EC EE ES FI GB GD GE GH GM HR HU ID IL IN IS JP  
KE KG KP KR KZ LC LK LR LS LT LU LV MA MD MG MK MN MW MX MZ  
NO NZ OM PH PL PT RO RU SD SE SG SI SK SL TJ TM TN TR TT TZ  
UA UG US UZ VN YU ZA ZM ZW  
RW (ARIPO): GH GM KE LS MW MZ SD SL SZ TZ UG ZM ZW  
RW (EAPO): AM AZ BY KG KZ MD RU TJ TM  
RW (EPO): AT BE BG CH CY CZ DE DK EE ES FI FR GB GR IE IT LU MC NL PT  
SE SK TR  
RW (OAPI): BF BJ CF CG CI CM GA GN GQ GW ML MR NE SN TD TG  
AI WO 2002-US34645 A 20021029  
PRAI US 2001-60/340,689 20011029

L3 ANSWER 4 OF 10 PCTFULL COPYRIGHT 2004 Univentio on STN  
AN 2002072017 PCTFULL ED 20020927 EW 200238  
TIEN METHOD OF DRUG TARGET VALIDATION  
TIFR PROCEDE DE VALIDATION DE CIBLES DE MEDICAMENTS  
IN SERAFINI, Tito, Andrew, 2249 Bunker Hill Drive, San Mateo, CA  
94402-3832, US [US, US]  
PA RENOVIS, INC., 270 Littlefield Avenue, South San Francisco, CA 94080, US  
[US, US], for all designates States except US;  
SERAFINI, Tito, Andrew, 2249 Bunker Hill Drive, San Mateo, CA  
94402-3832, US [US, US], for US only  
AG ANTLER, Adriane, M., Pennie & Edmonds LLP, 1155 Avenue of the Americas,  
New York, NY 10036, US  
LAF English  
LA English  
DT Patent  
PI WO 2002072017 A2 20020919  
DS W: AE AG AL AM AT AU AZ BA BB BG BR BY BZ CA CH CN CO CR CU CZ  
DE DK DM DZ EC EE ES FI GB GD GE GH GM HR HU ID IL IN IS JP  
KE KG KP KR KZ LC LK LR LS LT LU LV MA MD MG MK MN MW MX MZ  
NO NZ OM PH PL PT RO RU SD SE SG SI SK SL TJ TM TN TR TT TZ  
UA UG US UZ VN YU ZA ZM ZW  
RW (ARIPO): GH GM KE LS MW MZ SD SL SZ TZ UG ZM ZW  
RW (EAPO): AM AZ BY KG KZ MD RU TJ TM  
RW (EPO): AT BE CH CY DE DK ES FI FR GB GR IE IT LU MC NL PT SE TR  
RW (OAPI): BF BJ CF CG CI CM GA GN GQ GW ML MR NE SN TD TG  
AI WO 2002-US7294 A 20020312  
PRAI US 2001-60/275,073 20010312

L3 ANSWER 5 OF 10 PCTFULL COPYRIGHT 2004 Univentio on STN  
AN 2002064749 PCTFULL ED 20020904 EW 200234  
TIEN COLLECTIONS OF TRANSGENIC ANIMAL LINES (LIVING LIBRARY)  
TIFR COLLECTIONS DE LIGNEES D'ANIMAUX TRANSGENIQUES (BIBLIOTHEQUE VIVANTE)  
IN SERAFINI, Tito, Andrew, 2249 Bunker Hill Drive, San Mateo, CA  
94402-3832, US [US, US]  
PA RENOVIS, INC., 270 Littlefield Avenue, South San Francisco, CA 94080, US  
[US, US], for all designates States except US;  
SERAFINI, Tito, Andrew, 2249 Bunker Hill Drive, San Mateo, CA  
94402-3832, US [US, US], for US only  
AG ANTLER, Adriane, M., Pennie & Edmonds LLP, 1155 Avenue of the Americas,  
New York, NY 10036, US  
LAF English  
LA English  
DT Patent  
PI WO 2002064749 A2 20020822  
DS W: AE AG AL AM AT AU AZ BA BB BG BR BY BZ CA CH CN CO CR CU CZ  
DE DK DM DZ EC EE ES FI GB GD GE GH GM HR HU ID IL IN IS JP

**THIS PAGE BLANK (USPTO)**

KE KG KP KR KZ LC LK LR LS LT LU LV MA MD MG MK MN MW MX MZ  
NO NZ OM PH PL PT RO RU SD SE SG SI SK SL TJ TM TN TR TT TZ  
UA UG US UZ VN YU ZA ZM ZW

RW (ARIPO): GH GM KE LS MW MZ SD SL SZ TZ UG ZM ZW

RW (EAPO): AM AZ BY KG KZ MD RU TJ TM

RW (EPO): AT BE CH CY DE DK ES FI FR GB GR IE IT LU MC NL PT SE TR

RW (OAPI): BF BJ CF CG CI CM GA GN GQ GW ML MR NE SN TD TG

AI WO 2002-US4765 A 20020214  
PRAI US 2001-09/783,487 20010214

L3 ANSWER 6 OF 10 BIOSIS COPYRIGHT 2004 BIOLOGICAL ABSTRACTS INC. on STN  
DUPLICATE 1  
AN 2001:482536 BIOSIS  
DN PREV200100482536  
TI **Human preproneurotensin/neuromedin N.**  
AU Lal, Preeti [Inventor]; Shah, Purvi [Inventor]; Corley, Neil C. [Inventor,  
Reprint author]  
CS Mountain view, CA, USA  
ASSIGNEE: Incyte Genomics, Inc.  
PI US 6274720 August 14, 2001  
SO Official Gazette of the United States Patent and Trademark Office Patents,  
(Aug. 14, 2001) Vol. 1249, No. 2. e-file.  
CODEN: OGUPE7. ISSN: 0098-1133.  
DT Patent  
LA English  
ED Entered STN: 17 Oct 2001  
Last Updated on STN: 23 Feb 2002

L3 ANSWER 7 OF 10 PCTFULL COPYRIGHT 2004 Univentio on STN  
AN 2001083691 PCTFULL ED 20020826  
TIEN SYSTEM FOR IDENTIFYING AND ANALYZING EXPRESSION OF ARE-CONTAINING GENES  
TIFR SYSTEME D'IDENTIFICATION ET D'ANALYSE DE L'EXPRESSION DE GENES CONTENANT  
DES ELEMENTS RICHES EN ADENYLATE URIDYLATE (ARE)  
IN ABU-KHABAR, Khalid, S.;  
WILLIAMS, Bryan, R., G.;  
FREVEL, Mathias;  
SILVERMAN, Robert, H.  
PA THE CLEVELAND CLINIC FOUNDATION;  
KING FAISAL SPECIALIST HOSPITAL AND RESEARCH CENTRE;  
ABU-KHABAR, Khalid, S.;  
WILLIAMS, Bryan, R., G.;  
FREVEL, Mathias;  
SILVERMAN, Robert, H.  
DT Patent  
PI WO 2001083691 A2 20011108  
DS W: AE AG AL AM AT AU AZ BA BB BG BR BY BZ CA CH CN CR CU CZ DE  
DK DM DZ EE ES FI GB GD GE GH GM HR HU ID IL IN IS JP KE KG  
KP KR KZ LC LK LR LS LT LU LV MA MD MG MK MN MW MX MZ NO NZ  
PL PT RO RU SD SE SG SI SK SL TJ TM TR TT TZ UA UG US UZ VN  
YU ZA ZW GH GM KE LS MW MZ SD SL SZ TZ UG ZW AM AZ BY KG KZ  
MD RU TJ TM AT BE CH CY DE DK ES FI FR GB GR IE IT LU MC NL  
PT SE TR BF BJ CF CG CI CM GA GN GW ML MR NE SN TD TG  
AI WO 2001-US11993 A 20010412  
PRAI US 2000-60/196,870 20000412  
ICM G06F019-00  
ICS G01N033-483

L3 ANSWER 8 OF 10 BIOSIS COPYRIGHT 2004 BIOLOGICAL ABSTRACTS INC. on STN  
AN 2000:246144 BIOSIS  
DN PREV200000246144  
TI Comparative sequencing of the proneurotensin gene and association studies  
in schizophrenia.  
AU Austin, J.; Hoogendoorn, B.; Buckland, P.; Speight, G.; Cardno, A.; Bowen,  
T.; Williams, N.; Spurlock, G.; Sanders, R.; Jones, L.; Murphy, K.;  
McCarthy, G.; McGuffin, P.; Owen, M. J.; O'Donovan, M. C. [Reprint author]  
CS Division of Psychological Medicine, University of Wales College of  
Medicine, Heath Park, Cardiff, CF4 4XN, UK  
SO Molecular Psychiatry, (March, 2000) Vol. 5, No. 2, pp. 208-212. print.  
ISSN: 1359-4184.

**THIS PAGE BLANK (USPTO)**

DT Article  
LA English  
OS Genbank-S47339  
ED Entered STN: 14 Jun 2000  
Last Updated on STN: 5 Jan 2002

L3 ANSWER 9 OF 10 BIOSIS COPYRIGHT 2004 BIOLOGICAL ABSTRACTS INC. on STN  
AN 2000:339392 BIOSIS  
DN PREV200000339392  
TI Association analysis of the proneurotensin gene and bipolar disorder.  
AU Austin, J.; Hoogendoorn, B.; Buckland, P.; Jones, I.; McCandless, F.;  
Williams, N.; Middle, F.; Owen, M. J.; Craddock, N.; O'Donovan, M. C.  
[Reprint author]  
CS Division of Psychological Medicine, University of Wales College of  
Medicine, Heath Park, Cardiff, CF4 4XN, UK  
SO Psychiatric Genetics, (March, 2000) Vol. 10, No. 1, pp. 51-54. print.  
ISSN: 0955-8829.

DT Article  
LA English  
OS Genbank-S47339  
ED Entered STN: 10 Aug 2000  
Last Updated on STN: 7 Jan 2002

L3 ANSWER 10 OF 10 BIOSIS COPYRIGHT 2004 BIOLOGICAL ABSTRACTS INC. on STN  
AN 1999:187752 BIOSIS  
DN PREV199900187752  
TI Genotyping single nucleotide polymorphisms by primer extension and high  
performance liquid chromatography.  
AU Hoogendoorn, Bastiaan; Owen, Michael J.; Oefner, Peter J.; Williams,  
Nigel; Austin, Jehannine; O'Donovan, Michael C. [Reprint author]  
CS Division of Psychological Medicine, University of Wales College of  
Medicine, Heath Park, Cardiff, CF4 4XN, UK  
SO Human Genetics, (Jan., 1999) Vol. 104, No. 1, pp. 89-93. print.  
CODEN: HUGEDQ. ISSN: 0340-6717.

DT Article  
LA English  
ED Entered STN: 5 May 1999  
Last Updated on STN: 5 May 1999

**THIS PAGE BLANK (USPTO)**